

CRANFIELD UNIVERSITY

UMAIR SOORI

ENHANCED TARGET DETECTION IN CCTV NETWORK SYSTEM
USING COLOUR CONSTANCY

CRANFIELD DEFENCE AND SECURITY

PhD

Academic Year: 2010 - 2014

Supervisor: Peter Yuen

August 2014

CRANFIELD UNIVERSITY

CRANFIELD DEFENCE AND SECURITY

PhD

Academic Year 2010 - 2014

UMAIR SOORI

ENHANCED TARGET DETECTION IN CCTV NETWORK SYSTEM
USING COLOUR CONSTANCY

Supervisor: Peter Yuen
August 2014

© Cranfield University 2014. All rights reserved. No part of this publication may be reproduced without the written permission of the copyright owner.

ABSTRACT

The focus of this research is to study how targets can be more faithfully detected in a multi-camera CCTV network system using spectral feature for the detection. The objective of the work is to develop colour constancy (CC) methodology to help maintain the spectral feature of the scene into a constant stable state irrespective of variable illuminations and camera calibration issues.

Unlike previous work in the field of target detection, two versions of CC algorithms have been developed during the course of this work which are capable to maintain colour constancy for every image pixel in the scene: 1) a method termed as Enhanced Luminance Reflectance CC (ELRCC) which consists of a pixel-wise sigmoid function for an adaptive dynamic range compression, 2) Enhanced Target Detection and Recognition Colour Constancy (ETDCC) algorithm which employs a bidirectional pixel-wise non-linear transfer PWNLTF function, a centre-surround luminance enhancement and a Grey Edge white balancing routine.

The effectiveness of target detections for all developed CC algorithms have been validated using multi-camera 'Imagery Library for Intelligent Detection Systems' (iLIDS), 'Performance Evaluation of Tracking and Surveillance' (PETS) and 'Ground Truth Colour Chart' (GTCC) datasets. It is shown that the developed CC algorithms have enhanced target detection efficiency by over 175% compared with that without CC enhancement.

The contribution of this research has been one journal paper published in the Optical Engineering together with 3 conference papers in the subject of research.

Keywords:

Colour constancy; CCTV surveillance; Illumination invariance detection; Colour feature

LIST OF PUBLICATIONS

Journal Papers

Umair Soori, Peter Yuen, Ji Wen Han, Izzati Ibrahim, Wentao Chen, Kan Hong, Christian Merfort, David James and Mark Richardson, 2013. **Target recognitions in multiple-camera closed-circuit television using colour constancy**. *SPIE Optical Engineering*, 52(4), pp. 047202-1- 047202-13.

Izzati Ibrahim, Peter Yuen, Kan Hong, Tong Chen, Umair Soori, James Jackman and Mark Richardson, 2012. “**Illumination invariance and shadow compensation via spectro-polarimetry technique**”. *SPIE Optical Engineering*, 51(1), pp. 107004-1-107004-10.

Umair Soori, Peter Yuen, Ji Wen Han, Senthurran Selvagumar, Mengjia Ding, David James and Mark Richardson. **Enhanced Target Detection In CCTV Network System using Colour Constancy**. Paper to be submitted to the Int. Journal of Remote Sensing

Conference Papers

Umair Soori, PWT Yuen, J Han, D James and M Richardson “**Target Recognitions in Multiple Camera CCTV Using Colour Constancy**” 2013 Oral Presentation: Defence Academy of the United Kingdom. 11th Electro-optics & Infrared Conference Shrivenham, Swindon

Umair Soori, PWT Yuen, I Ibrahim, J Han, K Hong, T Chen, J Jackman, D James and M Richardson “**Colour Invariant Target Recognition in Multiple Camera CCTV Surveillance**” 2012. Poster Presentation: Defence Academy of the United Kingdom. 10th Electro-optics & Infrared Conference Shrivenham, Swindon.

Umair Soori, P. W. T. Yuen I. Ibrahim J. Han A. Tsitiridis K. Hong T. Chen J. Jackman D. James and M. Richardson, “**Colour invariant target recognition**

in multiple camera CCTV surveillance” 2011. SPIE Proceedings of Optics and Photonics for Counterterrorism and Crime Fighting VII.

T Chen, PWT Yuen , K Hong, I Ibrahim, A Tsitiridis, **Umair Soori**, J Jackman, D James and M Richardson, “***Assessment of tissue blood perfusion in-vitro using hyperspectral and thermal imaging techniques***” 2011. Bioinformatics and Biomedical Engineering, (iCBBE) 2011 5th International Conference.

Tsitiridis A, Yuen PWT, Ibrahim I, **Umair Soori**, Chen T, Hong K, James D and Richardson M, “***Enhanced object recognitions in cortex like machine vision***” 2011. Artificial Intelligence Applications and Innovations IFIP Advances in Information and Communication Technology.

ACKNOWLEDGEMENTS

I would like to express my perpetual gratitude to my supervisor Dr. Peter Yuen for his help, guidance, suggestion, ambition and trust throughout my research work. His thoughtful feedback turned this into a much better piece of work than it was. I do not have words to express my gratitude for his support and worthy suggestions. I would like to thank the MOD/CPNI, Cranfield University for the partial support of the project.

Foremost I would like to express my heartiest gratitude to my family, my caring wife Farah and my daughter Filza, and especially to my parents. I would dedicate this work to my family because without their countless emotional and moral support, guidance and understanding I would never have been able to get to this stage. I am also thankful to my sisters and brother for their encouragement and prayers.

I am particularly appreciative to my lab mates at Cranfield University they all made my work enjoyable, their cooperative nature made my work easier to accomplish on time.

TABLE OF CONTENTS

ABSTRACT	i
LIST OF PUBLICATIONS.....	ii
ACKNOWLEDGEMENTS.....	iv
LIST OF FIGURES.....	viii
LIST OF TABLES	xiv
LIST OF EQUATIONS.....	xv
LIST OF ABBREVIATIONS	xx
Chapter 1: Introduction	23
1.1 Problem Statement: CCTV surveillance technology.....	23
1.2 Issues for using spectral features in target detection	25
1.2 Objectives and Contributions	27
1.2.1 Overall Objectives	27
1.2.2 Contributions and Achievements.....	27
1.3 Organization of the thesis	28
Chapter 2: Overview of Related Work	29
2.1 Retinex Theory	29
2.1.1 Basic Retinex Methods.....	29
2.1.2 Advanced Retinex algorithms.....	32
2.2 White Balancing Methods	34
2.2.1 Reflection Model	34
2.2.2 Source Model	35
2.2.3 Static methods	35
2.2.4 Learning based methods.....	38
2.3 Person Re-identification techniques for surveillance.....	39
2.3.1 Brightness Transfer Function (BTF)	39
2.3.2 Textural Feature Based Detection.....	41
2.3.3 Colour Invariant Features.....	42
2.4 Mid tone dynamic compression	45
Chapter 3: Experimental Set Up and Dataset.....	49
3.1 Dataset	49
3.1.1 Imagery Library for Intelligent Detection System (i-LIDS) dataset....	49
3.1.2 PETS 2007 data set	50
3.1.3 GTCC dataset	51
3.2 Targets.....	53
3.3 Detectors	55
3.3.1 Anomaly Detection (AD).....	55
3.3.2 Matched Filter Detection	57
3.3.3 Parametric Classifier	58
3.3.4 Detector employed in this work: MD classifier	61
3.4 Experimental Set Up	62

3.4.1 Target Signature Acquisition	62
3.5 Assessment of detection performances	64
3.5.1 Figure of merit (FOM) using AUROC	64
3.5.2 Delta ΔE_{ab}	65
3.5.3 Colour Difference Percentage (CDP)	66
Chapter 4: Enhanced Luminance Reflectance Colour Constancy Algorithm (ELRCC) For CCTV Network	67
4.1 Colour Transfer (CT).....	67
4.2 Enhanced Luminance Reflectance Colour Constancy Algorithm (ELRCC)	68
4.2.1 Adaptive Dynamic Range Compression.....	69
4.3 Effects of V_{min} in Luminance Based CC Algorithm	71
4.3.1 Adaptive Mid Tone Frequency Components Enhancement	73
4.4 Results.....	76
4.4.1 Retinex and CT Results	76
4.4.2 ELRCC Results	81
4.4.3 Discussion.....	91
Chapter 5: Enhanced Target Detection Colour Constancy algorithm (ETDCC).....	93
5.1 Non-Linear Transfer Function (NLTF).....	95
5.2 Bidirectional Pixel Wise Non Linear Transfer Function	96
5.2.1 RGB to HSV	96
5.2.2 Pixel Wise Non Linear Transfer Function (PWNLTF).....	97
5.3 Centre Surround Luminance [CSL].....	102
5.4 Effectiveness of PWNLTF+CSL vs other NLTF algorithms.....	103
5.5 Grey Edge Algorithm.....	107
5.6 Colour invariant descriptors	108
5.7 Ultimate ETDCC formulation and test results	108
5.7.1 Targets Representations	108
5.7.2 Colour Descriptors.....	109
5.7.3 Comparison of Detectors.....	112
5.7.4 Detection Performances of Various CC Algorithms.....	114
5.7.5 ETDCC & other CC algorithms: detection of 10 targets	117
5.7.6 Discussion.....	121
Chapter 6: Experimental Discussion: All data sets [iLIDS, PETS, GTCC] ..	123
6.1 Targets.....	123
6.2 Limitation of colour invariant descriptors (rgb and C1C2C3).....	123
6.3 Colour constancy over multiple camera views	126
6.4 iLIDS & PETS data sets: All CC algorithms	127
6.4.1 Colour Enhancements.....	127
6.4.2 Colour Descriptors: EDTCC algorithm.....	131
6.4.3 Target detection results by ALL CC algorithms	136

6.5 GTCC data set: All CC algorithms	141
6.5.1 Colour enhancements	141
6.5.2 GTCC data set: ΔE_{ab} assessment.....	144
6.5.3 GTCC data set: ETDCC vs colour descriptors	147
Chapter 7: Conclusion and Future Works.....	149
7.1 Future works	151
REFERENCES.....	152
APPENDICES	163
Appendix A : Limitations of CC algorithms for target detections	163
Appendix B : Variable C in the PWNLTF	167
Appendix C : Frame to frame target detections	168

LIST OF FIGURES

- Figure 1-1: Shows the example of control room with large number of TV monitors, image is taken from (CCTV, (2014)). 23
- Figure 1-2: Highlight the effectiveness of using colour features for target detection: (top) greyscale image, (bottom) colour image. (image taken from (Gehler, *et al.* 2008)) 24
- Figure 1-3: To highlight one of the most concerning issues in machine vision for tracking targets from CCTV network footage: (a) raw image plot of target (T1) in camera 1 and the ROI of the target is depicted by the red box. (b) Exhibits the raw colour within the ROI for 10 frames of the video. (c) Same data as (b) but after the scene is transformed by ETDC algorithm developed in this work. (d) Depicts the Standard Deviation (SD) of the RGB attributes in the ROI for 10 frames of data before (in dots) and after processed by ETDC algorithm (triangle). 26
- Figure 3-1: (a) Illustrate schematic layout of camera view of the Gatwick Airport, (b) C1, C2, C3, C4 and C5 represents camera 1, camera 2, camera 3, camera 4 and camera 5 views respectively. 50
- Figure 3-2: Shows the camera view of the PETS 2007 dataset (PETS, 2007), C1, C2, C3 and C4 represent camera 1, camera 2, camera 3 and camera 4 views respectively. 51
- Figure 3-3: (a) shows ten variable illumination images of GTCC dataset from top left to bottom middle shows low to high illumination. (b) Illustrates the ground truth illumination values measured by light meter in lux with respect to (a). 53
- Figure 3-4: Shows the representative pictures of targets exploited in this study. (a) iLIDS dataset targets in camera 1. (b) PETS 2007 dataset targets in camera 1. (c) GTCC dataset 24 colour chart. 55
- Figure 3-5: Shows the detection ROC for target T3 in camera view #1 (denoted as T3C1) using the 3 detectors MD (circle), QD (triangle) and ACE (dot) for the raw data. All detection was performed using the colour opponent descriptor. 62
- Figure 3-6: Typical ROC results for the detection of a target in every 10 frames interval (Black plots) of a video clip when the target transverse across the scene. There are 10 detections for a particular target and each produces one ROC. (Red plot) for better visualise the detection performance for particular target within video clip, all ROC results as seen in (Black plots) is averaged into one (Red plot). 63
- Figure 3-7: Illustrates how the enhancement of target detection is assessed in this work: the figure of merit (FOM) for the detection enhancement is evaluated by the difference in areas of the two ROCs with respect to the raw data according to equation (3.24). The area is calculated within the

fixed PFA boundaries: in this case the enhancement is indicated by the area coloured in blue. Red trace: ROC of raw data, blue circle trace: ROC of the same data but after processed by ETDC. The detectors in both cases are the Mahalanobis Discriminant (MD) classifier. 65

Figure 4-1: Highlights one band intensity and luminance images. (a) Original/Raw image of Target (red box) in Cam 1. (b) Intensity image obtain by equation (4.2). (c) Luminance image of (a) obtain by equation (4.4) 69

Figure 4-2: Illustrates the effects of v_{min} settings on the colour constancy of a video sequence after processed by the luminance based algorithm. (a) Raw data, (b) $v_{min} = -5$ and (c) $v_{min} = -10$, (d) colour attributes of target (in red rectangle) after processing data using $v_{min} = -5$ and (e) colour attributes of target after processing data using $v_{min} = -10$ 72

Figure 4-3: Shows the effect of pixel wise dynamic range compression in the luminance based CC algorithm: (a) & (c) after processed by using a global mean v_{min} for the complete scene as proposed by (Tao, *et al.* 2005), (b) & (d) after processed by the proposed pixel wise v_{min} . (c) & (d) depict the colour characteristic of the target. Note that the images that have been processed by the global v_{min} in (a) is found not capable to rectify local colour non-uniformity due to variable illumination conditions..... 73

Figure 4-4: Outlines the ‘reflectance’ and one-band intensity map: (a) the intensity $I(x,y)$ of original image for Target (red box) in Cam1 (b) luminance based ‘reflectance’ image obtained by equation (4.3) (c) the one-band enhanced intensity I_{enh} by using equation (4.14). 75

Figure 4-5: Shows SSR result. (a) Raw/original Image with Target (red box) in cam 1. (b) SSR result with Gaussian surround constant “c” = 15. (c) SSR result with Gaussian surround constant “c” =80. (d) SSR result with Gaussian surround constant “c” = 250..... 77

Figure 4-6: Shows MSR result. (a) Raw/original Image with Target (red box) in cam 1. (b) MSR result with commonly used Gaussian surround constants $c_1=15$, $c_2=80$ and $c_3=250$ 77

Figure 4-7: Shows MSRCR result. (a) Raw/original Image with Target (red box) in cam 1..... 78

Figure 4-8: presents the raw and the processed data using SBD-MSR algorithm respectively in the left and right hand panels. In (a), (b) & (c) it shows the images of a target (red box) in clips extracted from Camera 1, 3 and 5 respectively..... 79

Figure 4-9: Illustrates the luminance (in right) of an image (left)..... 80

Figure 4-10: shows the raw (left) and processed data by LB MSR (right) for a target (red box) who had the appearance in (a) camera 1, (b) camera 3 and (c) camera 5. 81

Figure 4-11: Representative RGB pictures of 10 selected targets in this chapter: (left to right) T1 to T10, respectively. 82

Figure 4-12: Shows the averaged detection ROC of T3 ELRCC data in (a) camera view #1 (b) camera view #5. Note that some descriptors, such as the L1L2L3, have exhibited very poor colour invariance property thereby giving rather poor detection performances. The C1C2C3 descriptor also gives very variable results which can be seen from the ROC of the two camera views shown in (a) & (b). 83

Figure 4-13: Average of 10 detection results for target T3 in (a) camera view #1, (b) camera view #3, and (c) camera view #5, after the video sequence is processed by six different CC algorithms. Note that all CC algorithms, except for the one enhanced by the proposed ELRCC (in blue circle) method, exhibit little or no detection improvements with respect to that of the raw untreated data (in red cross) in some camera views..... 86

Figure 4-14: Shows representative RGB images selected from 3 camera views of target T10: (upper panel) raw data and (bottom panel) after processed by the ELRCC algorithm..... 88

Figure 4-15: Shows the mean of the ROC results for the detection of targets (a) T1, (b) T2, (c) T3, (d) T4 and (e) T5. The ROC of the raw data and that after processed by ELRCC are presented in circle and cross respectively. The three different camera views of data are presented in colour code of red, blue and green for view #1, view #2 and view #3 respectively. All detection was performed using the colour opponent descriptor. 91

Figure 4-16: Shows the averaged ROC detection results for all 10 targets in all camera views: the averaged raw ROC in red trace, and the average ROC for the ELRCC treated data in blue which shows a FOM of target enhancement by 90% FOM, implying for a PFA reduction by about a magnitude at PD=0.5..... 92

Figure 5-1: Outlines the flowchart of the proposed ETDC algorithm..... 94

Figure 5-2: Illustrates different non-linear transfer function plots. (a) AINDANE method with weighted parameter “z” (b) IRN transfer function curve with parameter v_{min} (c) NWIS curve with α and β parameters (d) PWNLTF curve with weighted parameter α 96

Figure 5-3: Outlines the bidirectional NLTF proposed in this work (a) first half of PWNLTF which is controlling dark part of images (b) Second half of PWNLTF which is controlling bright part of images. 98

Figure 5-4: Shows the effect of α and β (a) raw image (b) left: processed image with PWNLTF method right: its corresponding curve, where $\beta = 0.8$, similarly in (c) and (d) but this time $\beta = 0.9$ and $\beta = 1$ respectively. 101

Figure 5-5: Shows the comparison of PWNLTF with other NLTF algorithms on Target (red box) in camera 1, frame numbers 518 and 531 are the entry point of target where illumination is weak and frame numbers 596 and 622

are middle and exit point of target where illumination is strong. (a) to (e) are Raw, AINDANE, IRN, NWIS and PWNLTF+CSL respectively. 104

Figure 5-6: Highlights the RGB attributes of the ROI for 10 frames of Target (red box) in camera 1 after processed by PWNLTF+CSL and other NLTF algorithms as shown in Figure 5-5, (a) to (c) are R, G, and B bands respectively. Note that the images that have been processed by other NLTF algorithms are found not capable to rectify colour non-uniformity due to variable illumination conditions, (d) to (f) depicts the Standard Deviation (SD) of the R, G and B attributes respectively of the ROI for 10 frames, this SD gives the indication that inside the ROI colour losses its constancy as target moves in different illumination regions..... 105

Figure 5-7: Shows the effectiveness of CC for various algorithms in the colour opponent colour space. The circle and triangle plots represent data extracted from the first (ie frame #518) and last (frame#596) frame of images as shown in Figure 5-5..... 106

Figure 5-8 : Shows the representative pictures of targets exploited in this chapter: left to right T1 to T10 respectively..... 109

Figure 5-9: Shows the averaged detection ROC of T2 with respect various descriptor in (a) camera view #1 (b) camera view #3 and (c) camera view #5. 111

Figure 5-10: Shows the detection ROC for target T2 (a) camera view #1 (b) camera view #3 and (c) camera view #5 using the 3 detectors MD (circle), QD (square) and MF ACE (triangle) for the raw data and that after processed by CC algorithm (the ultimate ETDCC). 113

Figure 5-11: shows the average of 10 detection results for target T2 in (a) camera view #1, (b) camera view #3, (c) camera view #5, after the video sequence is processed by twelve different CC algorithms. Note that all CC algorithms, except that processed by the ETDCC (magenta star) method, exhibit little or no consistence detection improvements with respected to that of the raw untreated data (in red x)..... 116

Figure 5-12: Shows representative images selected from the all 3 camera views of target T1 (left to right camera 1, 2 and 3 respectively). (a) Raw data. (b) After transformed by the ultimate ETDCC algorithm..... 118

Figure 5-13: Shows the mean of the ROC results for the detection of targets (a) T1, (b) T2, (c) T4, (d) T6 and (e) T9. The ROC of the raw data and that after processed by ETDCC are presented in circle and cross respectively. The three different camera views of data are presented in colour code of red, blue and green for view #1, view #2 and view #3 respectively. Note that the circle plot ROC for all colour plots is substantially better than that of the cross in all cases. The colour opponent descriptor and the MD detector have been employed in all cases..... 120

- Figure 5-14: Shows the averaged ROC detection results for all 10 targets in all camera views, (red cross) ROC for raw data, (blue circle) ROC for after processing data with the ETDCC method. It can be seen that FOM of target enhancement by 155% FOM, implying for a PFA reduction by about two order of magnitude at PD=0.5..... 122
- Figure 6-1: Shows the representative pictures of targets exploited in this chapter (a) left to right T1 to T12 respectively extracted from iLIDS dataset (b) left to right T1 to T4 respectively extracted from PETS 2007 dataset..... 123
- Figure 6-2: Highlights the drawback of some colour invariant descriptors (a) raw RGB image extracted from GTCC dataset (b) after converted image into reduced rgb colour space (c) after converted image into C1C2C3 colour space..... 124
- Figure 6-3: shows another examples of limitations in some colour invariant features (a) raw RGB image from iLIDS (b) after converted image into rgb colour space (c) after converted into C1C2C3 colour space and (d) after processing the image using EDTCC..... 124
- Figure 6-4: Demonstrate the impacts of limitations in colour invariant descriptors on the target detection performance: (a) and (b) ROC for weak coloured target 9 in camera 1 and camera 3 respectively, (c) and (d) ROC for strong coloured target 7 in camera 1 and camera 3 respectively. 125
- Figure 6-5: Shows the O1 O2 O3 chrominance of target 1 for all three camera views after processed by all CC algorithms utilised in this study. The plots for views [1, 3, 5] are in [circle, triangle, square] and note that only the proposed ETDCC (in yellow) exhibit good colour constancy (i.e. small spread). 126
- Figure 6-6: Illustrates the RGB image of target 6 for 3 views from iLIDS dataset (a) raw images and (b)-(p) after processed by various CC algorithms. ... 131
- Figure 6-7: Illustrates the RGB image of 4 images from the GTCC dataset (a) raw images and (b)-(m) after processed by various CC algorithms..... 144
- Figure 6-8: Shows the mean intensity of each frame of image extracted from the GTCC data set after they are processed by various CC algorithms. The x-axis is the 10 experimental runs for recording the scene under 10 different irradiances. 144
- Figure A-1: Shows examples of targets that are difficult to detect using spectral feature: left to right FT1 (white), FT2 (grey) and FT3 (black). 164
- Figure A-2: Shows the ROC results for the detection of targets in 3 camera views of: (a) FT1 (white), (b) FT2 (grey) and (c) FT3 (black). 166
- Figure A-3: (a) top row: raw colour of FT2 (grey) for the 10 frames of the video clip. Bottom row: same data after transformed by the ETDCC algorithm. (b) Shows the ΔE_{ab} assessment of the data shown in (a). 166

Figure B-1: Show the various C values for the detection of T1 after the data is processed by PWNLTF using various C values. C=0.2 has been chosen for all the detection presented in this work..... 167

Figure C-1: Show an example of the frame-to-frame detection process employed in this work (a) Ground truth target map for the first 4 frames of images (b) ROC detections for the 10 frames of the raw data (c) ROC detections for the 10 frames of the data after treated by the ETDC algorithm, (d) Detection map of the first 4 frames of data using the raw data as test images at PFA 0.02, (e) Detection map of the first 4 frames of data after they are treated by the ETDC at PFA 0.02. 171

LIST OF TABLES

Table 4-1: Tabulates the Figure of Merit (FOM) in % for the detection of Target T3 in 3 different views after the data is treated by six different CC algorithms.....	87
Table 4-2: Tabulates the FOM for the detection performance enhancement for the data processed by the proposed ELRCC for all 10 targets.....	89
Table 5-1: Shows the efficiencies of different features utilised in this research study in FOM metric.....	111
Table 5-2: Tabulates the detection efficiencies of three different detectors that have been employed in this study.	114
Table 5-3: Tabulates the FOM for the detection of Target T2 in 3 different views after the data is treated by twelve different CC algorithms.	117
Table 5-4: Tabulates the FOM for the detection performance enhancement for the data processed by the proposed ETDC for all 10 targets.....	121
Table 6-1: Tabulates the FOM results for the detection of 12 targets in 3 different non-overlapping camera views of iLIDS data after the images treated by ETDC+ six different descriptors.	133
Table 6-2: Shows the FOM in % results for the detection of 4 targets in overlapping views of PETS data after images treated by ETDC+ six different descriptors.	135
Table 6-3: Tabulates the average FOM in % results for the detection of 16 targets by using ETDC + various colour descriptors.	135
Table 6-4: Compares the FOM in % for the detection of targets in the non-overlapping camera iLIDS data set after images processed by various CC algorithms.	138
Table 6-5: Compares the FOM in % for the detection of targets in the overlapping camera PETS data set after images processed by various CC algorithms.	140
Table 6-6: Shows the main result of this work which tabulates the FOM in % for the detection of targets in the iLIDS and PETS data set after images processed by various CC algorithms.	140
Table 6-7: Tabulates the ΔE_{ab} for processing the images in the GTCC data set by various CC algorithms. Each entry is the averaged ΔE_{ab} colour differences of 10 images recorded under step change of irradiance.	146
Table 6-8: Tabulates the CDP for processing the images in the GTCC data set by the ETDC algorithm together with various colour descriptors. Each entry is the averaged CDP over 10 images recorded under step change of irradiances.	148

LIST OF EQUATIONS

$Rr_i(x, y) = \log I_i(x, y) - \log(G(x, y) * I_i(x, y)) \dots (2.1)$	29
$I_i(x, y) = L_i(x, y)R_i(x, y) \dots (2.2)$	30
$R_{MSRi} = \sum_{n=1}^N \omega_n SSR \dots (2.3)$	30
$R_{MSRi} = \sum_{n=1}^N \omega_n \{ \log I_i(x, y) - \log[G(x, y) * I_i(x, y)] \} \dots (2.4)$	31
$R_{MSRCRi}(x, y) = C_i(x, y)R_{MSRi}(x, y) \dots (2.5)$	32
$C_i(x, y) = \beta \log(\alpha I_i'(x, y)) \dots (2.6)$	32
$I_i'(x, y) = I_i(x, y) / \sum_{i=1}^S I_i(x, y) \dots (2.7)$	32
$R_{MSRCRi}(x, y) = G[C_i(x, y) \sum_{n=1}^N \omega_n \{ \log I_i(x, y) - \log[G(x, y) * I_i(x, y)] \} + b] \dots (2.8)$	32
$\left\{ \begin{array}{l} \bar{R}_1 = R'_1, \quad n=1 \\ \bar{R}_n = R'_n - R'_{n-1}, \dots 2 \leq n \leq N \end{array} \right\} \dots (2.9)$	33
$R'_n = m \log(I) - m \log(F_n * I) \dots (2.10)$	33
$I(x, y) = R \frac{R}{R+G+B} + G \frac{G}{R+G+B} + B \frac{B}{R+G+B} \dots (2.11)$	33
$R_{MSR2}(x, y) = \frac{1}{3} \left(\log \frac{I(x, y).I(x, y).I(x, y)}{I(x, y) * (G_1(x, y) + G_2(x, y) + G_3(x, y))} \right) \dots (2.12)$	33
$R_j(x, y) = \sum_{j=1}^3 R_{MSR2}'(x, y) \left(\frac{I_j(x, y)}{I(x, y)} \right) \lambda \dots (2.13)$	34
$I_i(x, y) = \int_{\omega} R_i(x, y, \lambda) L(\lambda) S_i(\lambda) d\lambda \dots (2.14)$	35
$\begin{pmatrix} Rc \\ Gc \\ Bc \end{pmatrix} = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} Ru \\ Gu \\ Bu \end{pmatrix} \dots (2.15)$	35
$ke_c = \int f_c(\mathbf{x}) d\mathbf{x} \dots (2.16)$	36
$ke = \max f(\mathbf{x}) \dots (2.17)$	36
$ke = (\max R(\mathbf{x}), \max G(\mathbf{x}), \max B(\mathbf{x})) \dots (2.18)$	36
$e_c = \left(\int \left \frac{\partial^n I_c(\mathbf{x})}{\partial \mathbf{x}^n} \right ^p d\mathbf{x} \right)^{1/p} \dots (2.19)$	37
$I_i(\mathbf{x}) = m_b(\mathbf{x}) \int_{\omega} R_i(\mathbf{x}, \lambda) L(\lambda) S_i(\lambda) d\lambda + m_s(\mathbf{x}) \int_{\omega} L(\lambda) S_i(\lambda) d\lambda \dots (2.20)$	37
$\mathbf{M}_i I \in C \dots (2.21)$	38

$B_j = f_{ij}(B_i) \dots(2.22)$	40
$F_{ij}(B_i) = H_j^{-1}(H_i(B_i)) \dots(2.23)$	40
$MF_{ij} = \frac{1}{n} \sum_{i=1}^n f_i \dots(2.24)$	41
$FS_1 = 2B + (-R + G)$	
$FS_2 = 2B + (-G + R) \dots(2.25)$	44
$FS_3 = R + G + B$	
$L_1 = \frac{(R-G)^2}{(R-G)^2 + (R-B)^2 + (G-B)^2}$	
$L_2 = \frac{(R-B)^2}{(R-G)^2 + (R-B)^2 + (G-B)^2}$	44
$L_3 = \frac{(G-B)^2}{(R-G)^2 + (R-B)^2 + (G-B)^2}$	
$C_1 = \arctan\left(\frac{R}{\max\{G, B\}}\right)$	
$C_2 = \arctan\left(\frac{G}{\max\{R, B\}}\right) \dots(2.27)$	44
$C_3 = \arctan\left(\frac{B}{\max\{R, G\}}\right)$	
$r = R/(R + G + B)$	
$g = G/(R + G + B) \dots(2.28)$	45
$b = B/(R + G + B)$	
$O1 = \frac{R-G}{\sqrt{2}}$	
$O2 = \frac{R+G-(2xB)}{\sqrt{6}} \dots(2.29)$	45
$O3 = \frac{R+G+B}{\sqrt{3}}$	
$f(v) = \frac{1}{1+e^{-av}} \dots(2.30)$	46
$I'_n = \frac{I_n^{(0.75z+0.25)} + (1-I_n)0.4(1-z) + I_n^{(2-z)}}{2} \dots(2.31)$	46
$Z = \begin{cases} 0, \dots \text{for } I \leq 50 \\ \frac{I-50}{100}, \dots \text{for } 50 < I \leq 150 \\ 1, \dots \text{for } I > 150, \end{cases} \dots(2.32)$	46
$L'(x, y) = \frac{I(x, y) - 204}{51} I(x, y) + \left(1 - \frac{I(x, y) - 204}{51}\right) L(x, y) \dots(2.33)$	47
$L''_{enh}(x, y) = \frac{1}{1+e^{(-\alpha L)}} + \frac{1}{1+e^{(-\beta(L-10))}} - 0.5 \dots(2.34)$	47

$D_{RX}(\langle \mathbf{x} B \rangle) = [d(\mathbf{x}, B) > t] \dots (3.1)$	56
$\left[(\mathbf{x} - \hat{\mu})^T \hat{\Gamma}^{-1} (\mathbf{x} - \hat{\mu}) > t \right] \dots (3.2)$	56
$\left[\left\ \tilde{\mathbf{x}} \right\ ^2 \right] > t \dots (3.3)$	56
$\mu = \frac{1}{N} \sum_{n=1}^N \mathbf{x}(n) \dots (3.4)$	56
$\Gamma = \frac{1}{N} \sum_{n=1}^N (\mathbf{x}(n) - \mu)(\mathbf{x}(n) - \mu)^T \dots (3.5)$	56
$D_{AMF}(\mathbf{x} B, C) = \left[\frac{\tilde{\mathbf{s}}^T \tilde{\mathbf{x}}}{\ \tilde{\mathbf{s}}\ ^2} > t \right] \dots (3.6)$	57
$\left[\frac{(\mathbf{s} - \mu)^T \Gamma^{-1} (\mathbf{x} - \mu)}{(\mathbf{s} - \mu)^T \Gamma^{-1} (\mathbf{s} - \mu)} > t \right] \dots (3.7)$	57
$D_{ACE}(\mathbf{x} B, C) = \left[\frac{\tilde{\mathbf{x}}^T \tilde{\mathbf{P}}_s \tilde{\mathbf{x}}}{\tilde{\mathbf{x}}^T \tilde{\mathbf{x}}} > t \right] \dots (3.8)$	58
$[\cos^2 \alpha] > t \dots (3.9)$	58
$\tilde{\mathbf{P}}_s = \tilde{\Phi}_s (\tilde{\Phi}_s^T \tilde{\Phi}_s)^{-1} \tilde{\Phi}_s^T \dots (3.10)$	58
$\mathbf{x} \in w_i, \quad p(w_i \mathbf{x}) > p(w_j \mathbf{x}) \forall j \neq i \dots (3.11)$	58
$p(w_i \mathbf{x}) = p(\mathbf{x} w_i) p(w_i) / p(\mathbf{x}) \dots (3.12)$	59
$p(\mathbf{x}) = \sum_{i=1}^L p(\mathbf{x} w_i) p(w_i) \dots (3.13)$	59
$\mathbf{x} \in w_i, \quad p(\mathbf{x} w_i) p(w_i) > p(\mathbf{x} w_j) p(w_j) \forall j \neq i \dots (3.14)$	59
$g_i(\mathbf{x}) = \ln \{ p(\mathbf{x} w_i) p(w_i) \} = \ln p(\mathbf{x} w_i) + \ln p(w_i) \dots (3.15)$	59
$\mathbf{x} \in w_i, \quad g_i(\mathbf{x}) > g_j(\mathbf{x}) \forall j \neq i \dots (3.16)$	59
$p(\mathbf{x} w_i) = \frac{1}{(2\pi)^{N/2} \Gamma_i ^{1/2}} \exp \left\{ -\frac{1}{2} (\mathbf{x}_i - m_i)^T \Gamma_i^{-1} (\mathbf{x}_i - m_i) \right\} \dots (3.17)$	60
$g_i(\mathbf{x}) = \ln p(w_i) - \frac{N}{2} \ln(2\pi) - \frac{1}{2} \Gamma_i - \frac{1}{2} (\mathbf{x}_i - m_i)^T \Gamma_i^{-1} (\mathbf{x}_i - m_i) \dots (3.18)$	60
$g_i'(\mathbf{x}) = \Gamma_i + (\mathbf{x}_i - m_i)^T \Gamma_i^{-1} (\mathbf{x}_i - m_i) \dots (3.19)$	60
$\mathbf{x} \in w_i, \quad g_i'(\mathbf{x}) < g_j'(\mathbf{x}) \forall j \neq i \dots (3.20)$	60
$fd_i(\mathbf{x}) = (\mathbf{x}_i - m_i)^T \Gamma_i^{-1} (\mathbf{x}_i - m_i) \dots (3.21)$	61
$g_i(\mathbf{x}) = \ln p(w_i) - \frac{N}{2} \ln(2\pi) - \frac{1}{2} \sigma^{2N} - \frac{1}{2} (\mathbf{x}_i - m_i)^T \sigma^{-2} (\mathbf{x}_i - m_i) \dots (3.22)$	61
$g_i(\mathbf{x}) = -(\mathbf{x}_i - m_i)^T (\mathbf{x}_i - m_i) \dots (3.23)$	61

$$FOM = \left(\frac{AUROC_{cc} - AUROC_{raw}}{AUROC_{raw}} \right) 100 \quad (3.24) \dots\dots\dots 64$$

$$\Delta E_{ab} = \sum_{n=1}^{n=10} \sqrt{(L_n - L_{10th(max)})^2 + (a_n - a_{10th(max)})^2 + (b_n - b_{10th(max)})^2} \dots(3.25) \dots 65$$

$$CDP = \sum_{n=1}^{n=10} \sqrt{\left(\frac{B1_n - B1_{10th(max)}}{B1_{10th(max)}} \right)^2 + \left(\frac{B2_n - B2_{10th(max)}}{B2_{10th(max)}} \right)^2 + \left(\frac{B3_n - B3_{10th(max)}}{B3_{10th(max)}} \right)^2} \quad (3.26) \dots\dots 66$$

$$I'_{j(x,y)} = \mu_s + \left(\frac{\sigma_s}{\sigma_t} \right) (I_{j(x,y)} - \mu_t) \dots(4.1) \dots\dots\dots 68$$

$$I(x, y) = \max(r(x,y), g(x,y), b(x,y)) \dots(4.2) \dots\dots\dots 68$$

$$I(x, y) = L(x, y) R(x, y) \dots(4.3) \dots\dots\dots 68$$

$$L(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n) G(m+x, n+y) \dots(4.4) \dots\dots\dots 68$$

$$G(x, y) = q.e^{\left(\left(\frac{x^2+y^2}{c^2} \right) \right)} \dots(4.5) \dots\dots\dots 68$$

$$\iint q.e^{\left(\left(\frac{x^2+y^2}{c^2} \right) \right)} dx dy = 1 \dots(4.6) \dots\dots\dots 69$$

$$f(v) = \frac{1}{1+e^{-v}} \dots(4.7) \dots\dots\dots 70$$

$$L_n' = L_n [f(v_{max}) - f(v_{min})] + f(v_{min}) \dots(4.8) \dots\dots\dots 70$$

$$L_n'' = \ln\left(\frac{1}{L_n'} - 1 \right) \dots(4.9) \dots\dots\dots 70$$

$$L_{n,enh} = \frac{L_n'' - v_{min}}{v_{max} - v_{min}} \dots(4.10) \dots\dots\dots 70$$

$$v_{min} = \left(\left(\frac{I(x, y) - 1}{254} \right) k \right) - 6 \dots(4.11) \dots\dots\dots 70$$

$$L'_{n,enh}(x, y) = L_{n,enh}(x, y)^{E(x,y)} \dots(4.12) \dots\dots\dots 74$$

$$E(x, y) = \left(\frac{I_{conv}(x, y)}{I(x, y)} \right)^p \dots(4.13) \dots\dots\dots 74$$

$$I_{enh}(x, y) = L'_{n,enh}(x, y) R(x, y) \dots(4.14) \dots\dots\dots 74$$

$$ELRCC_{(R)}(x, y) = (I_{enh}(x, y) / I(x, y)) R(x, y)$$

$$ELRCC_{(G)}(x, y) = (I_{enh}(x, y) / I(x, y)) G(x, y) \dots(4.15) \dots\dots\dots 75$$

$$ELRCC_{(B)}(x, y) = (I_{enh}(x, y) / I(x, y)) B(x, y)$$

$$H = \left\{ \frac{1/2[(R-G) + (R-G)]}{(R-G)^2 + (R-B)(G-B)} \right\} \dots(5.1) \dots\dots\dots 96$$

where,

$$H = \begin{cases} H, & \text{if } B \leq G, \\ 360 - H, & \text{if } B > G \end{cases}$$

$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \dots (5.2)$	96
$V = \max(R, G, B) \dots (5.3)$	97
$I_{(x,y)} = T_{(x,y)}(x_{(x,y)}) \dots (5.4)$	97
$I_{(x,y)} = (x_{(x,y)})^T \dots (5.5)$	97
$I_{enh(x,y)} = \log(I_{(x,y)}) - \log(Z - I_{(x,y)}) \dots (5.6)$	97
$I_{enh(x,y)} = \alpha_{(x,y)}(\log(I_{(x,y)})) - (1 - \alpha_{(x,y)})(\log(Z - I_{(x,y)})) \dots (5.7)$	97
$\alpha_{(x,y)} = \left(\frac{I_{(x,y)}}{255} \right) (-c) + \beta \dots (5.8)$	97
$I_{enh}' = \frac{I_{enh} - I_{enh, \min}}{I_{enh, \max} - I_{enh, \min}} \dots (5.9)$	101
$HC_{(x,y)} = \frac{I_{raw} - I_{raw(avg)}}{I_{raw(avg)}}, HC_{(x,y)} = I_{enh}'_{avg} \dots (5.10)$	1022
$I_{conv}(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n)G(m+x, n+y) \dots (5.11)$	102
$G(x, y) = q \cdot e^{\left(-\left(\frac{x^2 + y^2}{c^2} \right) \right)} \dots (5.12)$	102
$I_{enh}''(x, y) = I_{enh}'(x, y)^{E(x,y)} \dots (5.13)$	102
$E(x, y) = \left(\frac{I_{conv}(x, y)}{I(x, y)} \right)^p \dots (5.14)$	103
$e_c = \left(\int \left \frac{\partial^n I_c(\mathbf{x})}{\partial \mathbf{x}^n} \right ^p d\mathbf{x} \right)^{1/p} \dots (5.15)$	107
$Ir_{ColourCorrection}(x) = \frac{1}{\sqrt{3}} \frac{Ir(x)}{e_r}$	
$Ig_{ColourCorrection}(x) = \frac{1}{\sqrt{3}} \frac{Ig(x)}{e_g} \dots (5.16)$	107
$Ib_{ColourCorrection}(x) = \frac{1}{\sqrt{3}} \frac{Ib(x)}{e_b}$	

LIST OF ABBREVIATIONS

ACE	Adaptive Coherence Estimator
AHE	Adaptive Histogram Equalisation
AINDANE	Adaptive And Integrated Neighbourhood Dependent Approach For Nonlinear Enhancement
AUROC	Area-Under ROC Assessments
BTF	Brightness Transfer Function
CBTF	Cumulative Brightness Transfer Function
CC	Colour Constancy
CCTV	Closed Circuit Tele-Vision
CDF	Cumulative Distribution Function
CDP	Colour Difference Percentage
CIE	Commission On Illumination
CNR	Contrast To Noise Ratio
CR	Colour Restoration
CSL	Centre Surround Luminance
CT	Colour Transfer
CWB	Customised White Balance
ELRCC	Enhanced Luminance Reflectance Colour Constancy Algorithm

ETDCC	Enhanced Target Detection And Recognition By Using Colour Constancy And Transfer Functions
FOM	Figure Of Merit
FOV	Field Of View
GTCC	Ground Truth Colour Chart
HDR	High dynamic range imaging
HOG	Histogram Of Oriented Gradients
HSV	Human Vision System
iLIDS	Imagery Library For Intelligent Detection Systems
IRN	Illuminance-Reflectance Nonlinear
LB-MSR	Luminance Based Multiple Scale Retinex
LIP	Logarithmic Image Processing
LPF	Low Pass Filter
MBTF	Mean Brightness Transfer Function
MD	Mahalanobis Discriminant
MF	Match Filter
MSCR	Maximally Stable Colour Regions
MSR	Multiple Scale Retinex
MSRCR	Multiple Scale Retinex With Colour Restoration
MWIS	Multilevel Windows Inverse Sigmoid

NLTF	Non-Linear Transfer Function
P re-ID	Person Re-Identification
PD	Probability Of Detection
PETS	Performance Evaluation Of Tracking And Surveillance
PFA	Probability Of False Alarm
PSNR	Peak Signal To Noise Ratio
PWNLTF	Pixel Wise Non-Linear Transfer Function
QD	Quadratic Discriminant
Retinex	Retina-And-Cortex System
RGB	Red Green Blue
RHSP	Recurrent High Structured Patches
ROC	Receiver Operating Characteristics
ROI	Region Of Interest
SBD-MSR	Sub-Band Decomposed Multiple Scale Retinex
SD	Standard Deviation
SDALF	Symmetry-Driven Accumulation Of Local Features
SSR	Single-Scale Retinex

Chapter 1: Introduction

1.1 Problem Statement: CCTV surveillance technology

One main objective of the Closed Circuit Tele-Vision (CCTV) surveillance network is the provision of visual recognitions and tracking of targets across wide areas (Soori, *et al.* 2011). Strategic places such as in the airports, banks, shopping malls, traffic control areas are commonly deployed CCTV network surveillance system for security reasons. It is reported that BBC News, 2006; McCahill & Norris, 2003) the number of operational CCTV (2014) cameras in the UK was over 4 million in 2002. However, live surveillance in CCTV network is in fact monitored manually by operator as depicted in Figure 1-1. This is highly ineffective as it is almost impossible for a person to monitor such large number of screens simultaneously (Smith, 2004).



Figure 1-1: Shows the example of control room with large number of TV monitors, image is taken from (CCTV, (2014)).

One solution to remedy this drawback is the development of an autonomous surveillance using machine vision technology. Autonomous target/people tracking in public places from CCTV footage have been a popular research since the turn of century. However, it is still a very active research topic despite of enormous effort has been put into the area over the last 20 years. The real difficulties for target detection in complex backgrounds are in fact due to

numerous factors, such as occlusions, viewing angles effect, variable illumination conditions and self-shadowing, which modulate intrinsic signature of targets thereby giving large false alarm rates in the detection. Face recognition technology in general cannot be applied in CCTV scenario because of the low resolution CCTV footages. Textural features are often corrupted in the crowded scene due to occlusions, and it is found that spectral features, such as the colour of people's dress and clothing may offer an alternative, and very often a more effective means for target detection particularly in the very crowded situation. This is highlighted in Figure 1-2 which depicts the image of a crowded scene in (top) greyscale and (bottom) colour. It is so obvious how easy to spot the various species of flowers just by their colours.

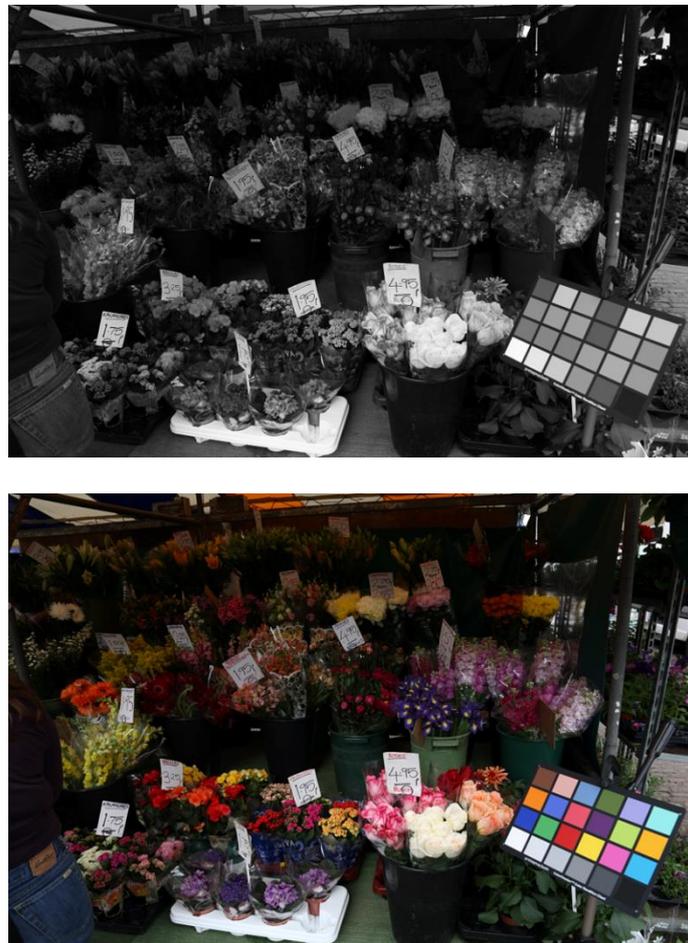


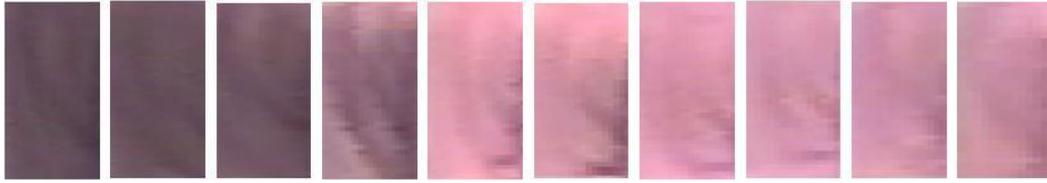
Figure 1-2: Highlight the effectiveness of using colour features for target detection: (top) greyscale image, (bottom) colour image. (image taken from (Gehler, *et al.* 2008))

1.2 Issues for using spectral features in target detection

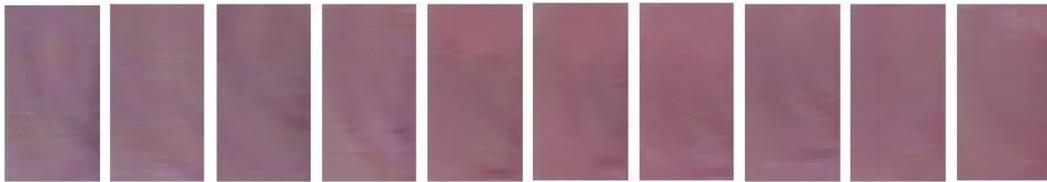
The spectral feature, or the colour of targets as seen by the sensor, is a compound effect of illumination conditions, background of the scene and the characteristic of the sensor. As illustrated in Figure 1-3 which highlights how the apparent colour of the target is changed when the scene is irradiated under variable illumination conditions. Figure 1-3(a) shows the raw RGB picture of a target in an airport scene and Figure 1-3(b) depicts the colour attributes of 10 frames of a subject as the target averaged over a small Region of Interest (ROI). It is seen from the plot in Figure 1-3(b) that the apparent shade and colour of the target seems to change rather dramatically when the target is walking towards the exit where the illumination is a lot stronger than other parts of the room. Figure 1-3(c) shows the colour attribute of the target after the image data is transformed by ETDCC algorithm developed in this PhD programme (see Chapter 5:). Unlike the raw data, it is seen that the shade and colour of the target remain to be more or less constant throughout the whole video sequence. Figure 3(d) plots the Standard Deviation (SD) of Figure 1-3(b) and Figure 1-3(c), which shows the effectiveness of the colour constancy to maintain the colour attributes of the target despite of the variable illumination conditions.



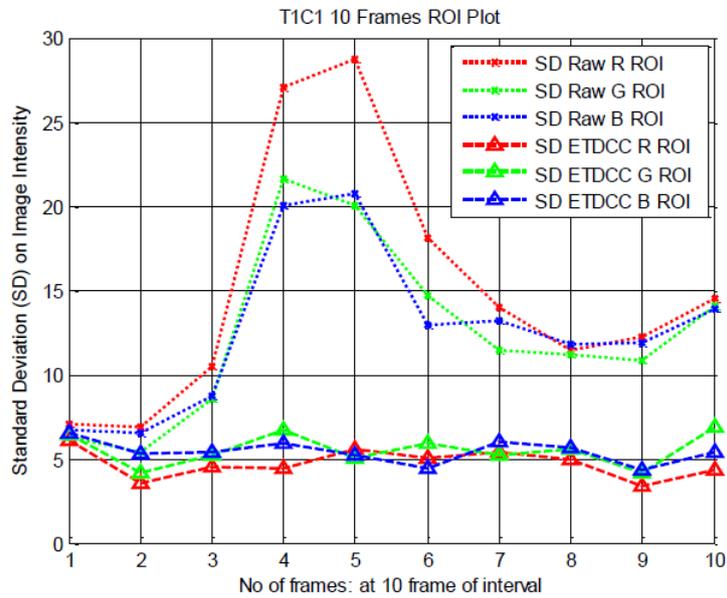
(a)



(b)



(c)



(d)

Figure 1-3: To highlight one of the most concerning issues in machine vision for tracking targets from CCTV network footage: (a) raw image plot of target (T1) in camera 1 and the ROI of the target is depicted by the red box. (b) Exhibits the raw colour within the ROI for 10 frames of the video. (c) Same data as (b) but after the scene is transformed by ETDCC algorithm developed in this work. (d) Depicts the Standard Deviation (SD) of the RGB attributes in the ROI for 10 frames of data before (in dots) and after processed by ETDCC algorithm (triangle).

1.2 Objectives and Contributions

1.2.1 Overall Objectives

The overall objectives of this research programme have been set as the following:

1. The development of colour constancy algorithm for broad band RGB imaging and to study its effectiveness for target detection purposes.
2. To assess the accuracy of detection of the targets in non-overlapping multiple camera views CCTV network by using ONLY colour as the detection feature, and to verify if colour constancy (CC) approach may help to reduce illumination artefacts and to improve target detection performance.
3. To assess the effectiveness of colour descriptors for maintaining colour constancy and target detection.

1.2.2 Contributions and Achievements

The main contributions of this PhD programme have been:

1. The development of a colour transfer mechanism to reduce the colour bleaching side effect in Retinex based CC algorithm.
2. The development of a luminance based CC algorithm which incorporated a pixel-wise mid-tone compression routine to further reduce colour bleach. This Enhanced Luminance Reflectance Colour Constancy algorithm (ELRCC) improves the target detection by over 100% compare to untreated raw data within the metric of Area-Under ROC assessments (AUROC). This work has resulted in a paper publication in the Optical Engineering Journal (Soori, *et al.* 2013).
3. Further development of ELRCC using a bidirectional pixel-wise non-linear transfer PWNLTF function, together with an in-scene multi-camera calibration method, a centre-surround luminance enhancement and a colour invariant descriptor to formulate an Enhanced Target Detection Colour Constancy (ETDCC) algorithm. This ETDCC achieves over 200% of target detection performances

better than that of using untreated raw data in the FOM metric. This work is being written up for another journal paper publication.

1.3 Organization of the thesis

Chapter One provides a brief introduction of the presented work. **Chapter Two** discusses the related work. **Chapter Three** elaborates about experimental set up and dataset utilised throughout this research programme. **Chapter Four** deals with two proposed modifications of luminance-based CC algorithms: one with a colour transfer mechanism and the other ELRCC that have been developed during the course of this work. **Chapter Five** details the Enhanced Target Detection Colour Constancy (ETDCC) algorithm which consists of a Pixel Wise Non-Linear Transfer Function (PWNLTF), a centre surround luminance enhancement and colour invariant descriptor. The target detection performance of this algorithm has been compared with several other CC work using 3 data sets in this chapter. **Chapter Six** presents the experimental results and discussions and the thesis is concluded together with an outlook of future work in **Chapter Seven**.

Chapter 2: Overview of Related Work

2.1 Retinex Theory

Retinex theory (RETINa-and-cortEX system) was firstly proposed by E. Land and McCann (1971) to model the lightness and colour perception of the human vision system. Land studied the characteristic of lightness and colour observation of human eye and derived a theory to implement visual perception for machine vision. Retinex is based on the surround spatially opponent operation which is related to the neurophysiological functions of neurons in the retina and cerebral cortex in the primates visual perception. Retinex has been one of the earliest CC methods for coping colour perception in non-uniform illumination scenarios.

Retinex theory can be divided into two different hypotheses: firstly, it provides a model of human vision based CC (Meyer, 2010) which will be discussed in more detail in section 2.2.3. Secondly it also gives basic concept for image enhancement (Jobson, *et al.* 1997b; Rahman, *et al.* 1996; Barnard & Funt, 1997) which will be discussed here.

2.1.1 Basic Retinex Methods

2.1.1.1 Single-scale Retinex (SSR)

Jobson *et al.* (1997b); (1996) refines the Retinex model and proposed a single-scale Retinex (SSR), which applies a fixed scale (filter) for all colour channels of the image. The SSR can either provide dynamic range compression by utilising small scale or tonal rendition by using large scale. In this methodology the illumination is firstly estimated through a Gaussian low pass filter (LPF) to an input image. The output image is then obtained by subtracting the logarithmic input image to logarithmic Gaussian filtered image. The Single-scale Retinex is given by equation (2.1)

$$Rr_i(x, y) = \log I_i(x, y) - \log(G(x, y) * I_i(x, y)) \quad (2.1)$$

Where $Rr_i(x, y)$ is the Retinex output, $I_i(x, y)$ is image distribution in the i -th colour band in (R,G,B) channel, “*” denotes the convolution operation and

$G_i(x, y)$ is the normalised surround Gaussian function given by $G(x, y) = ke^{-(x^2+y^2)/c^2}$ where c is the Gaussian surround space constant, and k is a normalisation constant which satisfy $\iint G(x, y)dxdy = 1$. The image is the product of scenes reflectance and illumination as shown in equation (2.2)

$$I_i(x, y) = L_i(x, y)R_i(x, y) \quad (2.2)$$

where $L_i(x, y)$ is the spatial distribution of illumination and $R_i(x, y)$ the distribution of scene reflectance. Jobson highlighted that Gaussian filter has the ability to have more regional effect and that it can give better dynamic range compression through different space constants. However, one drawback of this methodology is that the space constant c is a user defined empirical variable which cannot be computed mathematically.

2.1.1.2 Multiscale Retinex (MSR)

Hurlbert (1989) extended Land's (1971) Retinex theory from one surround constant into three different sigma values of the Gaussian surround function in order to obtain better dynamic range compressions for each colour channel. Hurlbert's work has been the first of its kind which leads to other Retinex derivatives (Rahman, *et al.* 1996; Barnard & Funt, 1997; Tao & Asari, 2003), such as the Multi Scale Retinex (MSR) (Rahman, *et al.* 1997; Jobson, *et al.* 1997a) which is formulated by a weighted combination of multi-scale components of 3 colour SSR. The use of multiple scales in each colour channels in MSR has the advantage to maintain a better balance between dynamic range compression and colour rendition. The MSR formulation is shown in equations (2.3) and (2.4):

$$R_{MSRi} = \sum_{n=1}^N \omega_n SSR \quad (2.3)$$

$$R_{MSRi} = \sum_{n=1}^N \omega_n \{ \log I_i(x, y) - \log [G(x, y) * I_i(x, y)] \} \quad (2.4)$$

where $i \in (R, G, B)$, N is the number of the scales and ω_n is the weighting factor of each scale under constraint of summation $(\omega_n)=1$, R_{MSRi} is the i -th colour channel of the linearly combined SSR. It seems that systematic analysis of how to choose the optimum number of scales and the exact weighting values in each colour channel are still lacking. Most researchers in the field have employed three scales for the image analysis, and it is difficult to judge whether their analysis is meaningful. The weights have direct effects of emphasising more on dynamic range compression or colour rendition. Chao (2007) has presented a rather detailed analysis using SSR and MSR on MRI images. A wide range of scales and weights have been employed in many MSR work in the past, and the results have shown that Retinex methods indeed outperformed the histogram equalisation based methods. Zhang, *et al.* (2011) also concluded that when 'good' parameterisation of MSR is attained it not only maintaining good contrast ratio but also that it increases the entropy (visual quality) of image in comparison to histogram equalisation and SSR methods.

2.1.1.3 Multiple Scale Retinex with Colour Restoration (MSRCR)

MSR is basically formulated from SSR, which, implements the centre and surround by subtracting off the logarithmic of the surround (or logarithmic ratio) thereby inducing artefacts such as colour bleaching particularly when the space constants are not set properly. This centre and surround side effect imposes great drawbacks in both MSR and SSR, which tends to turn the image into 'grey'. This is very unfavourable to target detection especially when colour feature is used for the detection. To help reduce this colour bleaching issue, a Colour Restoration (CR) algorithm has been proposed known as Multi-Scale Retinex with Colour Restoration (MSRCR) (Barnard & Funt, 1997; Rahman, *et al.* 1998; Rahman, *et al.* 1997; Jobson, *et al.* 1997a). The colour restoration essentially estimates the statistics of the colour attributes from the raw image and it is then transferred into colour constancy processed image to improve the

colour integrity. The CR and the sum of multi-scales in MSRCR can induce side effect of out of the range which can be constraint by compressing the histogram of the outputs to within the range through an auto gain/offset mechanism. The MSRCR is formulated as in equation (2.5):

$$R_{MSRCRi}(x, y) = C_i(x, y)R_{MSRi}(x, y) \quad (2.5)$$

Several linear and nonlinear functions have been implemented and Jobson (1997b); (1996) found that the general colour restoration is much better in the form of equation (2.6)

$$C_i(x, y) = \beta \log(\alpha I_i'(x, y)) \quad (2.6)$$

$$I_i'(x, y) = I_i(x, y) / \sum_{i=1}^S I_i(x, y) \quad (2.7)$$

$$R_{MSRCRi}(x, y) = G[C_i(x, y) \sum_{n=1}^N \omega_n \{\log I_i(x, y) - \log[G(x, y) * I_i(x, y)]\} + b] \quad (2.8)$$

Where α is strength of non-linearity and β is control gain constant, $I_i(x, y)$ is the raw image. The G and b are the gain and offset and all these parameters are user defined function (see section 4.4.1).

2.1.2 Advanced Retinex algorithms

2.1.2.1 Sub Band Decomposition MSR (SBD-MSR)

Sub-band decomposition MSR (SBD-MSR) (Jang, *et al.* 2008) basically is the MSR with enhancements of having an additional two-state transfer functions for mid-tone dynamic compression, then follow by a procedure for decomposing scale components to make them more independent to each other. The author (Jang, *et al.* 2008) nomenclatures this centre surrounds operation and scale component decomposition as *mlog*. The algorithm firstly maps the MSR output through a two state logarithmic function which effectively moderate the colour attributes in both extreme ends of the intensity attributes. Then a differential

scale component is applied to reduce the overlaps between scales as shown in equations (2.9) and 2.10):

$$\begin{cases} \overline{R}_1 = R'_1, & n = 1 \\ \overline{R}_n = R'_n - R'_{n-1}, \dots, 2 \leq n \leq N \end{cases} \quad (2.9)$$

$$R'_n = m \log(I) - m \log(F_n * I) \quad (2.10)$$

where R'_n and \overline{R}_n denotes the moderated Retinex output and the sub-band decomposed (differential component) Retinex output (SD-retinex output) respectively. N is the total number of scales, G_n is the Gaussian space function and I is the input image. In the standard MSR, a user-defined constant gain is needed for the calculation of the Retinex output, but in SD-MSR the gain for each sub-band can be evaluated from the differential component.

2.1.2.2 Luminance Based MSR (LBMSR)

Luminance Based MSR (LB-MSR) (Tao & Asari, 2003; Sun, *et al.* 2007) implements the centre and surround function using the luminance of each pixel, instead of manipulating the colour attributes of the image as in the conventional MSR/SSR method. The luminance can be obtained through PCA transform (Sun, *et al.* 2008) or as defined by (Tao & Asari, 2003):

$$I(x, y) = R \frac{R}{R+G+B} + G \frac{G}{R+G+B} + B \frac{B}{R+G+B} \quad (2.11)$$

Where I is the intensity map. In this methodology the LB-MSR only requires one cycle of centre surround convolution, instead of three cycles that are required in the MSR/SSR methodology thus reduces computational cost significantly. The centre and surround in LBMSR method is then reduced into:

$$R_{MSR2}(x, y) = \frac{1}{3} \left(\log \frac{I(x, y).I(x, y).I(x, y)}{I(x, y) * (G_1(x, y) + G_2(x, y) + G_3(x, y))} \right) \quad (2.12)$$

Where (G_1, G_2, G_3) are the Gaussian space function for the 3 colour channels, $I(x,y)$ is the intensity map of the image. This is the main difference between the LBMSR with respected to the conventional MSR/SSR.

Similar to the MSR the gain-offset is applied on the convoluted R_{MSR2} and the output is mapped into (0-255) range such that the zero point is set by the minimum intensity of the scene. Again the parameters of the gain (the contrast) and offset (brightness) are scene dependent and are set by user. To compensate colour bleach in LBMSR, a colour transfer through the chromaticity of the original image can be incorporated as shown in equation (2.13)

$$R_j(x, y) = \sum_{j=1}^3 R_{MSR2}'(x, y) \left(\frac{I_j(x, y)}{I(x, y)} \right) \cdot \lambda \quad (2.13)$$

Where $j = 1$: Red component, $j = 2$: Green component, $j = 3$: Blue component, λ = Flexible adaptive constant between 0 and 1.

2.2 White Balancing Methods

The ability of Retinex to maintain Colour Constancy (CC) is largely based upon the centre and surrounds to deduce the luminance of the scene and to ratio (or logarithmic subtraction) with the sensor (raw) data. This is similar to white balancing, which resolves to find the spectral irradiance (illumination) and subsequently correction of the estimated illumination such that the resulted image appears to be taken under white light or canonical light source. These methods commonly assume uniform illuminations across the scene. Note that image formation not only depending upon the intrinsic reflectivity of targets, but also that factors such as irradiance, sensor characteristics and background adjacency effects are also important too.

2.2.1 Reflection Model

Reflection model defines the physical interaction of scene with light and the simplest one is the Lambert reflectance model which defined as equation (2.14).

$$I_i(x, y) = \int_{\omega} R_i(x, y, \lambda) L(\lambda) S_i(\lambda) d\lambda \quad (2.14)$$

Where $R_i(x, y, \lambda)$ donates the surface reflectance of targets in the scene, $L(\lambda)$ is the spectral irradiance and $S_i(\lambda)$ is the sensor spectral response/characteristic, $i = [R, G, B]$. All of these variables are defined as a function of the wavelength (λ) over spectral region ω . $I_i(x, y)$ is the corresponding intensity map of the i -th colour channel. In general if the sensor property is known then the variation of the colour in an image is directly dependent on the target reflectivity and the illumination source.

2.2.2 Source Model

One method to maintain CC is to transform the unknown illumination light source of the scene into a known light source. Most of the transformation is linear especially in the broad band [RGB] imaging (Kries, 1970) and it rescales the unknown source to match with the known. The von kries model is given in equation (2.15):

$$\begin{pmatrix} Rc \\ Gc \\ Bc \end{pmatrix} = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix} \begin{pmatrix} Ru \\ Gu \\ Bu \end{pmatrix} \quad (2.15)$$

Where $[Ru, Gu, Bu]^T$ is the RGB channels of input image taken under unknown light source, $[Rc, Gc, Bc]^T$ is the RGB channels of output image under conical light source and the diagonal matrix is the scale factors which maps the colour of unknown light source in u to that of known in c .

2.2.3 Static methods

Colour Constancy (CC) in general can be roughly divided into two main classes of i) static method where no prior information was needed, ii) learning based methods, where training or learning is required prior to analysis. Apart from Retinex, Grey-World (Buchsbaum, 1980) has been a common technique in the

static approach, which postulates the average colour of a scene is achromatic when it is illuminated by a canonical light source. This means that the mean value of the R, G, B channels of an image would be a common grey value and any deviation from grey is due to the variation of light source. The image f can be defined as (equation (2.16)):

$$ke_c = \int f_c(\mathbf{x})d\mathbf{x} \quad (2.16)$$

where “e” is the colour of light source, k is the constant $k=0$ for no reflectance and $k=1$ for total reflectance, c is the number of channels ($c = R, G, B$) and \mathbf{x} is the coordinate in the image.

The white patch method (Land, 1977) which is also known as the perfect reflectance method, postulates that the highest intensity of an image is due to the surface colour of the target in the scene. The Max-RGB algorithm processes each colour channel separately (Land, 1977):

$$ke = \max f(\mathbf{x}) \quad (2.17)$$

$$ke = (\max R(\mathbf{x}), \max G(\mathbf{x}), \max B(\mathbf{x})) \quad (2.18)$$

where “e” is the colour of illuminate source, k is the reflectance, c is the number of channels ($c = R, G, B$) and \mathbf{x} is the coordinate of the image. Finlayson & Trezzi (2004) proposed a new method using shades of grey and have employed Minokwski norms for source estimation. They showed that Max-RGB method is equivalent to $P=\infty$ Minkowski norm and the Grey World is equivalent to $P=1$ Minkowski norm. The method shades of grey was shown to outperform the Grey World and Max-RGB methods when the dataset of Ciurea & Funt (2003) is used. They also showed that better CC result can be achieved by using $P=6$ Norm. Another technique of this type of method is Grey Edge (Weijer, *et al.* 2007) which postulates that the average reflectance difference in the scene is achromatic. Unlike the above methods Grey Edge is based on the derivatives of the whole scene. The average of the colour derivatives can be used to compute

the light source colour. Their study has shown that the distribution of colour derivatives can reveal the differences in the illumination directions of the source.

$$e_c = \left(\int \left| \frac{\partial^n I_c(\mathbf{x})}{\partial \mathbf{x}^n} \right|^P d\mathbf{x} \right)^{1/P} \quad (2.19)$$

where e is estimated light colour, c is the number of band ($c = R, G, B$), P denote the Minkowski norm and n is Grey-Edge order. If $n = 0$ and for $P = 1$ then the equation (2.19) is equal to the Grey-World assumption. For $P = \infty$ it is equal to White-Patch algorithm and for Shades of Grey $1 < P < \infty$. Finally for higher n (that is $n > 1$) it is similar to the Grey-Edge method (Weijer, *et al.* 2007). The advancement of illumination estimation algorithm for CC was presented in (Gijsenij, *et al.* 2012). Instead of using the complete scene the author employed local image patches, which was found more suitable to situations in which multiple light sources were present in the scene. This method has shown a better detection performance than the Grey world particularly when the scene is illuminated by several different light sources.

Other static methods, such as the dichromatic reflection model as shown in equation (2.20) have also been proposed (Finlayson & Schaefer, 2001b). Dichromatic reflection model utilises physical intersections between the objects and irradiance of the scene. It is postulated that the image planes of a scene due to different wavelength of illuminations are statistically different as the result of the different reflective characteristics of objects in different spectral regions across the scene. Thus the intersection between the planes from different spectral channels can be used to estimate colour of light source.

$$I_i(\mathbf{x}) = m_b(\mathbf{x}) \int_{\omega} R_i(\mathbf{x}, \lambda) L(\lambda) S_i(\lambda) d\lambda + m_s(\mathbf{x}) \int_{\omega} L(\lambda) S_i(\lambda) d\lambda \quad (2.20)$$

where image values $I_i(\mathbf{x})$ depend upon the colour of illuminant $L(\lambda)$, surface reflectance $R_i(\mathbf{x}, \lambda)$ and the sensor sensitivity $S_i(\lambda)$. All of these variables are defined as a function of the wavelength (λ) over visible spectrum w and i -th

colour channel ($i = R, G, B$). The $m_b(\mathbf{x})$ and $m_s(\mathbf{x})$ are the scale parameters for diffuse and specular reflection respectively.

Other workers in the field (Lee, 1986; Klinker, *et al.* 1988; Finlayson & Schaefer, 2001a) make use of technique that estimates the colour of illumination by comparing chromaticity of two surfaces. The estimation is achieved by searching the intersection between two or more dichromatic lines in the chromaticity colour space. Tan, *et al.* (2004) employed inverse-intensity chromaticity space to recover specular reflections which is then used for restoring the colour of the light source.

2.2.4 Learning based methods

2.2.4.1 Gamut based Method

Gamut mapping is a training based CC method first proposed by Forsyth (1990). The algorithm assumes that only limited number of colours is significant in a real world image for a given known light source, thereby any change of colour that appears in an image is due to the change of illuminant source. The set of colours that should have existed in a known light source can be learned offline from real world datasets. The learning from the ground truth illumination is labelled as canonical gamut C . The input gamut I is generated from the input image taken under unknown light source it is then mapped to match with that of the canonical gamut C through the diagonal matrix \mathbf{M}_i :

$$\mathbf{M}_i I \in C \quad (2.21)$$

Early gamut method makes use of RGB colour space which is also known as 3D gamut mapping algorithm. Finlayson (1996) introduced an extension of Forsyth's work (1990) in which a 2D chromaticity space for gamut mapping has been employed. Barnard (1997) proposed the modification of diagonal model by introducing a feasible mapping function in the 2D chromaticity colour space and the 3D RGB colour space. Gijsenij, *et al.* (2010) extends gamut mapping by incorporating higher order statistics of images. Moreover, it is suggested that a

more robust result can be achieved by considering the intersection of multiple feasible sets of the mapping.

2.2.4.2 Machine Learning Based Methods

Machine learning based methods deduce the colour of light sources by using a model which is based on the supervised training from ground truth dataset. Training using Neural Network (NN) methodology (Cardei, *et al.* 2002; Funt & Cardei, 1999) such as Multi-Layer Perceptron (MLP) feed-forward neural network in chromaticity colour spaces has shown a better result than other colour constancy algorithms. Funt & Carde (1999) utilised bootstrapping method which synthetically generate large amount of images from small amount of real dataset and have achieved good results.

Bayesian theory has also been employed in which the probability distributions of the illuminant and reflectance are learned. After the prior probability is computed from the training dataset, the illuminant is estimated from the probability distribution under the assumption that the data is Gaussian. The posterior can be estimated using Maximum a Posteriori (MAP) or Minimum Mean Squared Error (MMSE). Freeman & Brainard (1995) proposed a maximum local mass (MLM) and has shown better CC performance than that of the Grey-World, subspace technique and the MAP estimator. Rosenberg, *et al.* (2003) developed a Bayesian approach for CC using non- nonparametric Gaussian models with performance better than gamut mapping.

2.3 Person Re-identification techniques for surveillance

This section reviews common techniques for person re-identification which also known as people tracking or people recognition especially in a multi-camera network for surveillance applications.

2.3.1 Brightness Transfer Function (BTF)

Although research in person re-identification (P re-ID) has received considerable attentions over the past few decades (Porikli, 2003; Javed, *et al.* 2005; Javed, *et al.* 2008; Prosser, *et al.* 2008), people tracking is still one of the

hot topics within the machine vision community nowadays. One significant issue in the application of P re-ID is the inconsistency of image brightness across the camera network which alleviates the rate of false alarm and subsequently fails the target detection.

To address this network inconsistency issue, Javed *et al.* (2005); (2008) and Porikli (2003) proposed a Brightness Transfer Function (BTF) to model the brightness between different camera views in order to compensate inconsistent illuminations across the entire scene. This algorithm assumes a constant illumination over each camera view. The BTF f_{ij} is a function for a pair of camera views C_i and C_j and a given pair of targets O_i and O_j . This BTF maps brightness value B_i in O_i to the corresponding brightness value B_j in O_j as shown in equation (2.22).

$$B_j = f_{ij}(B_i) \quad (2.22)$$

The f_{ij} ideally should be operated in pixel domain. However, limitations due to non-rigidity of target's shape and size as the result of geometric factors such as occlusions and different field of views (FOV) normalised histograms of the whole object or even for the entire scene has been the common approach to exploit this technique in practise. It is assumed that the percentage of image points in an observation O_i with brightness less than or equal to B_i is equal to the percentage of image points in the observation O_j with brightness less than or equal to B_j (Javed, *et al.* 2005). Given H_i and H_j to be the normalised cumulative histograms for O_i and O_j observations respectively, then the BTF can be computed as (equation (2.23))

$$F_{ij}(B_i) = H_j^{-1}(H_i(B_i)) \quad (2.23)$$

where H_j^{-1} is the inverted cumulative histogram. Note that $H_i(B_i)$ and $H_j(B_j)$ are readily available from the pair of the scene data. The F_{ij} is simply the transfer function between the two scenes. To improve the accuracy the Mean

Brightness Transfer Function (MBTF) is often employed by averaging all objects in views C_i and C_j . For n targets in two camera views the BTF f_{ij} can be computed for each target then $M = (f_1, f_2, f_3, \dots, f_n)$ and computed as (equation (2.24))

$$MF_{ij} = \frac{1}{n} \sum_{i=1}^n f_i \quad (2.24)$$

Alternative approach using the Cumulative Brightness Transfer Function (CBTF) has been reported (Prosser, *et al.* 2008). The CBTF is computed by accumulating all target observations in one view and to evaluate the BTF from this cumulative histogram. Prosser, *et al.* (2008) employed this approach together with a bidirectional matching scheme to reduce false positive in their people tracking work. Some authors also suggested using an adaptive training scheme to overcome local illumination variation (Chen, *et al.* 2008). One drawback in this approach is that the BTF is only applicable in the overlapped camera view network system.

2.3.2 Textural Feature Based Detection

Textural features have been commonly employed for target detection over half a century. Scale invariant SURF/SIFT features together with KD-tree approach for fast matching have been reported for people tracking applications (Hamdoun, *et al.* 2008). Gray & Tao (2008) proposed viewpoint invariant target recognition by combining colour and texture features to improve the tracking (Gray, *et al.* 2007). Multiple colour spaces of [RGB, YCBCR, HSV] and texture features of Schmid and Gabor filters have been employed. The authors have shown that better tracking can be achieved by using more weight in the spectral (colour) feature.

Gheissari, *et al.* (2006) has employed spatiotemporal technique to obtain appearance invariant feature through the colour and structural clues of people's clothing. Bak, *et al.* (2010) has also employed appearance based technique to detect target's body parts such as head, limbs and torso by using colour and

texture approach. The body parts are recognised by using Histogram of Oriented Gradients (HOG) with covariance based descriptor which is used to find similarity over the target and the scene. It is found that the colour feature extraction is not straight forward even for a single camera view. The authors have resolved to use histogram equalisation on each RGB bands for the colour feature extraction. Note that histogram equalisation is not a robust CC technique (see section 6.4.3).

Further examples of spectral textural processing have been the Symmetry-Driven Accumulation of Local Features (SDALF) technique which makes use of three robust features: 1) colour histogram in HSV colour space, 2) Maximally Stable Colour Regions (MSCR) used for colour region displacement measurement and 3) Recurrent High Structured Patches (RHSP) for target texture classification (Farenzena, *et al.* 2010). Others workers (Gray & Tao, 2008; Wang, *et al.* 2007) have also utilised colour, shape and appearance model for target detection. Their results have shown that features extracted from the RGB and Lab colour space in general perform better than that of the HSV colour space. Gilbert & Bowden (2006) have proposed an incremental online learning methodology to model the incremental colour variations and the probability distributions of spatio-temporal links of targets amongst multiple cameras. The proposed method reported an increase of detection accuracy over time without prior information of the scene/target. Note that the target appearance methodology is more prone to errors particularly when occlusion is present in the scene.

2.3.3 Colour Invariant Features

Direct use of colour attributes in machine vision target detection is not favourable due to the many factors that can moderate the intensities of the pixel image. Many authors (Orwell, *et al.* 1999; Mittal & Davis, 2003; Park, *et al.* 2006) have opted to use colour or chrominance feature for target detection. These colour features have been exploited in various colour spaces such as YCBCR/HSV/RGB (Orwell, *et al.* 1999; Mittal & Davis 2003; Park, *et al.* 2006).

The hue (H component) has been found more effective than the saturation (S) and value (V) components for people tracking in HSV space.

Swain & Ballard (1990) has achieved good people tracking using a colour histogram intersection method which works well in strong occlusion and background interference situations. The detection is invariant to viewing angles but it is not robust enough for scenes with variable illuminations across the network. Funt & Finlayson (1995) have improved Swain's model by using illumination invariant colour derivative feature which is similar to that of the centre-surround cells in human's vision. The algorithm was tested under 3 different colour temperatures and achieved almost 100% of target detection comparing to the ~75% detection rate using Swain's model. Gevers and Smeulders (1999) extended the work of Funt and Finlayson (1995) by proposing several colour invariant descriptors such as the C1C2C3 and L1L2L3 and the authors have validated their invariance properties with respect to the theory. The performance of these colour descriptor for tracking people has been tested and compared with other CC algorithms (Soori, *et al.*, 2013). It has been found that although these colour descriptors maintain colour constancy rather well, the people tracking ability of these descriptors are not as good as other CC models.

While these chrominance features maintains CC rather well by removing the pixel intensity, the chrominance in general somehow reduces the prominence of feature making it not as acute as that in the original colour space. For instance, high intensity images with good signal to noise ratio contain more information for target detection than that in the dark areas. Thus by removing the intensity features will result in a loss of valuable information or clues. Moreover chrominance cannot distinguish different shade of grey (see section 6.2).

There are many colour invariant features reported in the literature and here is a brief overview of the most common colour descriptors reported in the literature.

2.3.3.1 Sum Feature

The sum feature is commonly exploited in machine vision (Mehetre, *et al.* 1995) and (Soori, *et al.*, 2011) and it is defined as shown in equation (2.25)

$$\begin{aligned}FS_1 &= 2B + (-R + G) \\FS_2 &= 2B + (-G + R) \\FS_3 &= R + G + B\end{aligned}\tag{2.25}$$

2.3.3.2 L1 L2 L3 Feature

The L1L2L3 feature has the property to define the direction of the triangular colour plane in the *RGB*-space (Gevers & Smeulders, 1999):

$$\begin{aligned}L_1 &= \frac{(R-G)^2}{(R-G)^2 + (R-B)^2 + (G-B)^2} \\L_2 &= \frac{(R-B)^2}{(R-G)^2 + (R-B)^2 + (G-B)^2} \\L_3 &= \frac{(G-B)^2}{(R-G)^2 + (R-B)^2 + (G-B)^2}\end{aligned}\tag{2.26}$$

2.3.3.3 C1 C2 C3 Feature

C1C2C3 feature makes use of colour ratio and arctan function which has been considered as one of the robust colour descriptors for object detection (Gevers & Smeulders, 1999):

$$\begin{aligned}C_1 &= \arctan\left(\frac{R}{\max\{G, B\}}\right) \\C_2 &= \arctan\left(\frac{G}{\max\{R, B\}}\right) \\C_3 &= \arctan\left(\frac{B}{\max\{R, G\}}\right)\end{aligned}\tag{2.27}$$

2.3.3.4 rgb Colour Feature

The rgb colour descriptor is also known as normalised RGB colour feature as it involves the normalisation of the pixel intensity (Gevers & Smeulders, 1999):

$$\begin{aligned}r &= R/(R + G + B) \\g &= G/(R + G + B) \\b &= B/(R + G + B)\end{aligned}\tag{2.28}$$

2.3.3.5 Colour Opponent Feature

The colour opponent (CO) (Gijsenij & Gevers, 2011) has been commonly exploited in machine vision as it provides better colour invariance. The colour opponent feature is given in equation (2.29)

$$\begin{aligned}O1 &= \frac{R - G}{\sqrt{2}} \\O2 &= \frac{R + G - (2 \times B)}{\sqrt{6}} \\O3 &= \frac{R + G + B}{\sqrt{3}}\end{aligned}\tag{2.29}$$

where R, G and B are the Red, Green and Blue bands of an RGB image respectively. O1 is commonly known as Red-Green opponency, O2 is known as Blue-Yellow opponency and O3 is the average of all bands.

2.4 Mid tone dynamic compression

High dynamic range imaging (HDR) is the technique to improve the quality of image via a non-linear transfer of the input intensity into the limited output range. The most common non-linear transfer function (NLTF) has been the logarithmic one which has been deployed in almost all of the display devices such as TV, computer monitors, photographic and image displays.

Apart from the logarithmic compression, a number of non-linear transfer functions have been proposed to help improve colour constancy in the machine vision research area (Tao, *et al.* 2005; Tao & Asari, 2005; Tao, *et al.*, 2006). Tao & Asari (2005) has suggested a sigmoid function (equation (2.30))

$$f(v) = \frac{1}{1 + e^{-av}} \quad (2.30)$$

where “a” is constant and v is a parameter that can be evaluated from global mean of the image intensity map I_m . Soori *et al.* (2013) has refined the transfer function by implementing a more effective pixel-wise v_{\min} parameter for better colour corrections. The implementation of pixel-wise v_{\min} will be discussed in section 4.3 in more detail. Tao & Asari (2005) have proposed an Adaptive and Integrated Neighbourhood Dependent Approach for Nonlinear Enhancement (AINDANE) (Tao & Asari, 2005) which employs cumulative distribution function (CDF) of the intensity image for an adaptive contrast enhancement. AINDANE improves the visual quality of image especially for images recorded under insufficient illumination. AINDANE transfer function is in the form of:

$$I'_n = \frac{I_n^{(0.75z+0.25)} + (1-I_n)0.4(1-z) + I_n^{(2-z)}}{2} \quad (2.31)$$

where I_n is the input intensity of the pixel at (x,y), I'_n output intensity transformed by AINDANE transfer function, Z is the image dependent parameter:

$$Z = \begin{cases} 0, \dots \text{for } I \leq 50 \\ \frac{I-50}{100}, \dots \text{for } 50 < I \leq 150 \\ 1, \dots \text{for } I > 150, \end{cases} \quad (2.32)$$

where I is the intensity level at 0.1 of the cumulative distribution function (CDF) of the whole scene. Note that the CDF is the normalised cumulative histogram of the entire scene, and therefore this algorithm cannot address issue such as local self-shadowing. Ghimire & Lee (2011); (2012) extend the concept by

dividing the image into sub-sets of overlapping blocks and image dependent parameter z can be computed from each block. Asari *et al.* (2006) has proposed a Multilevel Windows Inverse Sigmoid (MWIS) function (Asari, *et al.* 2006) which is designed for extreme non-uniform illumination scenarios. This method employs histogram and global mean dependant parameters to control the exact form of the transfer function. In this algorithm the illumination and intensity values is averaged first (as shown in equation (2.33)) to reduce halo effect:

$$L'(x, y) = \frac{I(x, y) - 204}{51} I(x, y) + \left(1 - \frac{I(x, y) - 204}{51}\right) L(x, y) \quad (2.33)$$

where I is input intensity image and L is illuminance of input image. After averaging L' it is then normalised into L'' and the mid tone compression is in the form of (equation (2.34)):

$$L''_{enh}(x, y) = \frac{1}{1 + e^{(-\alpha L')}} + \frac{1}{1 + e^{(-\beta(L' - 10))}} - 0.5 \quad (2.34)$$

where parameter α , β are the control parameters for dark and bright regions respectively. The control parameters can be assessed in-scene through the global mean or through user-defined threshold values. Other form of the compression such as the cosine based non-linear transfer function (Arigela & Asari, 2006) has shown better CC performance as compared to the bare MSRRCR, sigmoid function (Tao & Asari, 2005) and AINDANE (Tao & Asari, 2005). Wavelet based function together with hyperbolic sine function has been suggested by Unaldi *et al.* (2009). The results further confirm the advantage of non-linear compression with respected to the linear colour restoration procedure like that of colour restoration (CR) in MSRRCR (see section 2.1.1.3). Other forms of dynamic compressions, such as that based on hyperbolic tangent Tsai & Chou., 2011; Tsai, 2012), arcsine function (Arigela & Asari, 2013), logarithmic (Choudhury & Medioni, 2009; Choudhury & Medioni, 2010; Jourlin & Pinoli, 1987; Jourlin & Pinoli, 1988) and parameterised logarithmic (Panetta, *et al.*, 2011) have been proposed and most of them produce better image enhancement and CC performance with respect to that without NLTF.

It is noted that the proposed dynamic compression techniques are all suffering from drawbacks of:

1. The control parameters of the transfer function are mostly estimated from the entire scene which cannot handle local effects such as self-shadowing.
2. Most of the proposed algorithms address the enhancement of image quality in the dark regions. It is necessary to deal with compression for over-exposed area too.

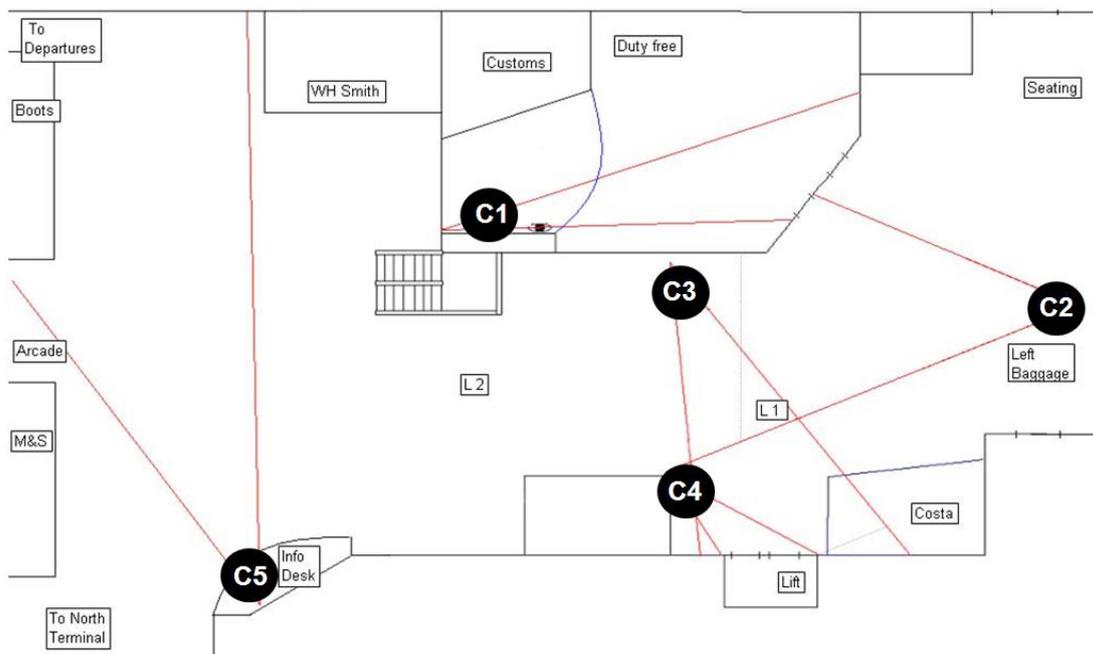
Chapter 3: Experimental Set Up and Dataset

3.1 Dataset

Three different type of datasets: 1) Imagery Library for Intelligent Detection System (i-LIDS) dataset (UK Home Office, 2008), 2) Performance Evaluation of Tracking and Surveillance (PETS) dataset (PETS, 2007) and 3) Ground Truth Colour Chart (GTCC) dataset have been employed throughout this study.

3.1.1 Imagery Library for Intelligent Detection System (i-LIDS) dataset

i-LIDS dataset (UK Home Office, 2008) consists of serials of CCTV video footage recorded in the Gatwick airport. The data is purchased from the Home Office Scientific Branch. The footage covers five different camera views that are non-overlapping. The data set has a total of ~44 hours of video, consisting of 107 shots of 5-camera extracts from 12 different collection epochs. Figure 3-1 shows the schematic layout of camera views and locations. The data presented in this study is selected from one of the series MCT-TR (1001-1005).



(a)



(b)

Figure 3-1: (a) Illustrate schematic layout of camera view of the Gatwick Airport, (b) C1, C2, C3, C4 and C5 represents camera 1, camera 2, camera 3, camera 4 and camera 5 views respectively.

3.1.2 PETS 2007 data set

PETS (2007) dataset which is obtained online from (PETS, 2007) consists of four overlapping camera views. PETS (2007) dataset is mainly designed for security/criminal events, within a real-world environment. The scenarios are recorded from four cameras and involve multiple actors. Namely 3 situations are 1) loitering, 2) attended luggage removal (theft) and 3) unattended luggage. Moreover equidistant markers were also placed on the floor of the terminal for the purposes of calibration. However we do not utilised any prior information in our experiments and colour (spectral) feature has been solely used for the target detection. Figure 3-2 shows the example of four different overlapping camera views. Note here that the camera view 3 is the imaging through a glass window which makes the colour of the scene appears to be quite different. This also makes target detections quite challenging in this case.

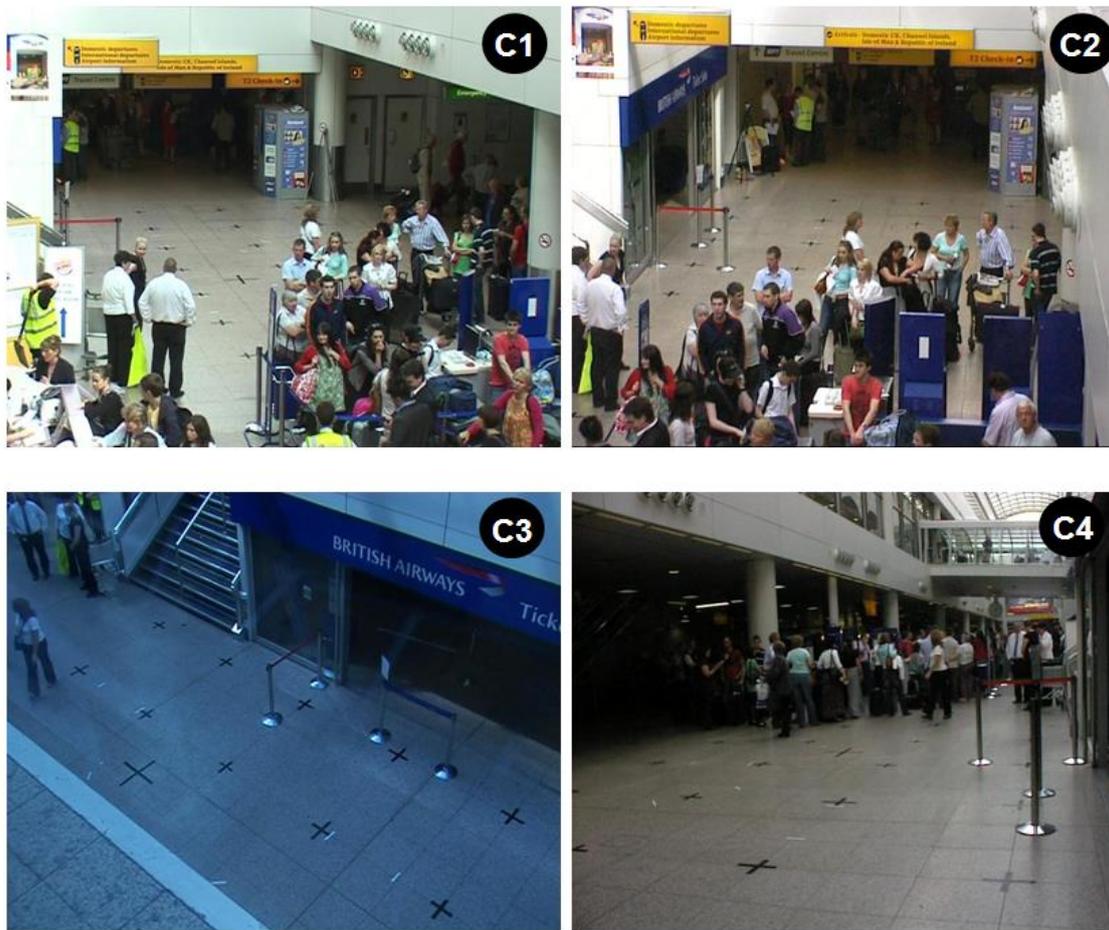
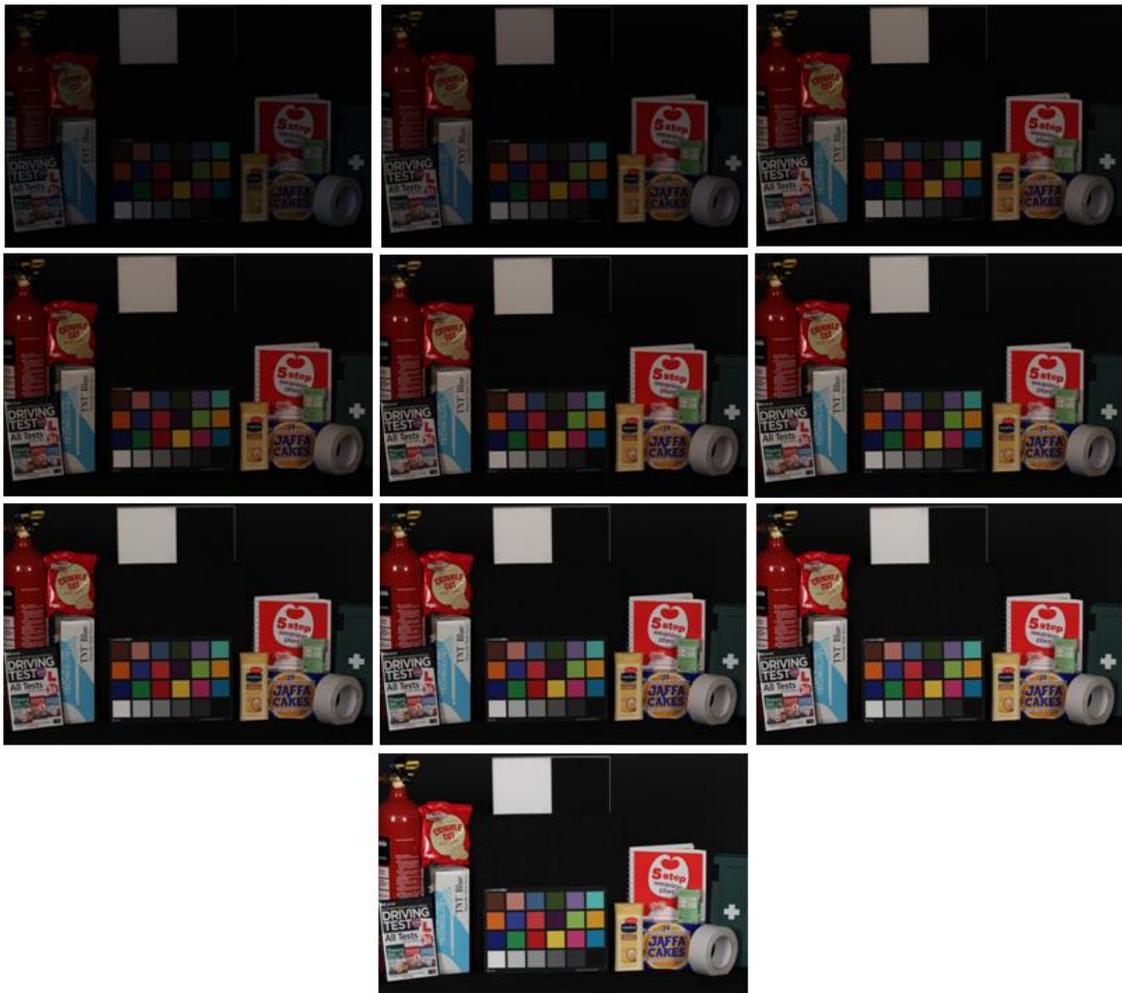


Figure 3-2: Shows the camera view of the PETS 2007 dataset (PETS, 2007), C1, C2, C3 and C4 represent camera 1, camera 2, camera 3 and camera 4 views respectively.

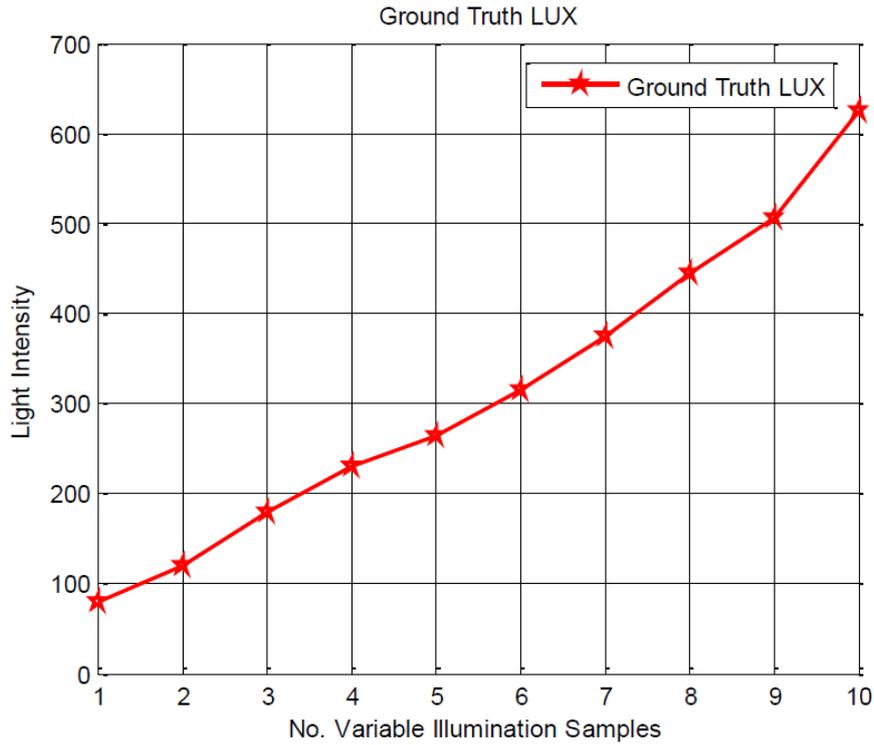
3.1.3 GTCC dataset

GTCC dataset is recorded by us in the laboratory using Canon 600D SLR Camera with ground truth data. In this data set the Gretag MacBeth Colour Checker Chart has been used as target which consists of 24 different colour boxes with various colour substances in the background. The scene is irradiated by two 1000W Halogen lamps with diffuser box at about 3m away from the target. The ground truth is obtained by setting the camera in customised white balance (CWB) mode for taking images of the white/black spectralon as references. During the experiment the camera is set in manual

mode with Aperture = F5.6, Shutter Speed = 1/15, ISO = 100 and MF manual focussing. Images are taken at 10 different step changes of irradiance and the illumination in each step is measured by light meter. The raw images of the 10 illumination settings are shown in Figure 3-3(a). Figure 3-3(b) illustrates the ground truth lux values with respect to images depicted in Figure 3-3(a)



(a)



(b)

Figure 3-3: (a) shows ten variable illumination images of GTCC dataset from top left to bottom middle shows low to high illumination. (b) Illustrates the ground truth illumination values measured by light meter in lux with respect to (a).

3.2 Targets

A number of human and colour patch targets have been chosen from the 3 data sets (see fig Figure 3-4):

1. iLIDS data set: 28 people targets in various colour and grey clothing.
2. PETS data set: 5 people targets in various colour and grey clothing.
3. GTCC data set: 24 various colour and grey patches.



(a)



(b)



(c)

Figure 3-4: Shows the representative pictures of targets exploited in this study. (a) iLIDS dataset targets in camera 1. (b) PETS 2007 dataset targets in camera 1. (c) GTCC dataset 24 colour chart.

3.3 Detectors

3.3.1 Anomaly Detection (AD)

Anomaly detection is a technique that locates and identifies uncommon pixel vector with reference to a background model. If an observed pixel spectra

deviates from the ‘norm’ of the background, this deviation is measured in a distance metric to represent the degree of anomaly. The background characteristics can be estimate from the scene, and hence AD is a detector that does not need the prior knowledge of targets or background (Ahlberg & Renhorn, 2004).

AD algorithm is firstly developed by Irving Reed and Xiaoli Yu in 1990 and the algorithm is commonly known as RX detector. Yu and Reed (1990) developed the algorithm under the generalized likelihood ratio test (GLRT) framework for multidimensional image data assuming that the spectrum of the target and the covariance of the background are in Gaussian distributions and they are generally unknown.

By denoting a background model as B , a distance measure as $d(\cdot)$ and a threshold is t , pixel x is regard as anomaly if (Reed & Xiaoli, 1990; Stein, *et al.* 2002).

$$D_{RX}(\langle \mathbf{x} | B \rangle) = [d(\mathbf{x}, B) > t] \quad (3.1)$$

$$= \left[(\mathbf{x} - \hat{\mu})^T \hat{\Gamma}^{-1} (\mathbf{x} - \hat{\mu}) > t \right] \quad (3.2)$$

$$= \left[\left\| \tilde{x} \right\|^2 \right] > t \quad (3.3)$$

Where D_{RX} is the RX detector, μ is the mean and Γ covariance matrix which can be computed as follows:

$$\mu = \frac{1}{N} \sum_{n=1}^N \mathbf{x}(n) \quad (3.4)$$

$$\Gamma = \frac{1}{N} \sum_{n=1}^N (\mathbf{x}(n) - \mu)(\mathbf{x}(n) - \mu)^T \quad (3.5)$$

In equation (3.5), the distance between each pixel signature to the background signature is calculated using Mahalanobis distance. Here, we assume that the background is model as a multivariate Gaussian distribution.

Hence in AD it is the model of the background which is needed to identify the uncommon pixels. There are various techniques reported for characterizing the background such as the employment of local or global windowing, target removal or clustering.

3.3.2 Matched Filter Detection

Matched filter detection is a well-developed technique using a known target spectral signatures or target probe to search for the presence of that spectrum in a scene (West, 2005). It attempts to detect and locate pixels containing a target material of known spectral composition. In here, the scene background remained as unstructured background or stochastic background which takes the form of first and second order spectral statistics (mean and covariance) estimated from the scene data like that done in the anomaly detection.

3.3.2.1 The adaptive matched filter (AMF)

The adaptive matched filter (AMF) is a detector that models and suppresses an unstructured background and then uses a known target spectrum, denoted as \mathbf{s} , to search for that in the scene. The pixel \mathbf{x} is classify as target classes, C if

$$D_{AMF}(\mathbf{x} | B, C) = \left[\frac{\tilde{\mathbf{s}}^T \tilde{\mathbf{x}}}{\|\tilde{\mathbf{s}}\|^2} \right] > t \quad (3.6)$$

$$= \left[\frac{(\mathbf{s} - \boldsymbol{\mu})^T \boldsymbol{\Gamma}^{-1} (\mathbf{x} - \boldsymbol{\mu})}{(\mathbf{s} - \boldsymbol{\mu})^T \boldsymbol{\Gamma}^{-1} (\mathbf{s} - \boldsymbol{\mu})} \right] > t \quad (3.7)$$

This detector is optimum only when the target and background follow Gaussian distribution, and in real applications it is highly unlikely (Ahlberg & Renhorn, 2004).

3.3.2.2 Adaptive coherence/cosine detector (ACE)

The adaptive coherence or cosine detectors model the target variability by measuring the angle between the target pixels to the target spectrum. The pixel \mathbf{x} is classified as a target pixel if (Ahlberg & Renhorn, 2004; Manolakis, *et al.*, 2003).

$$D_{ACE}(\mathbf{x} | B, C) = \left[\frac{\tilde{\mathbf{x}}^T \tilde{\mathbf{P}}_s \tilde{\mathbf{x}}}{\tilde{\mathbf{x}}^T \tilde{\mathbf{x}}} \right] > t \quad (3.8)$$

$$= [\cos^2 \alpha] > t \quad (3.9)$$

Where $\tilde{\mathbf{P}}_s$ is the projection and reconstruction operator onto the d target subspace, that is

$$\tilde{\mathbf{P}}_s = \tilde{\Phi}_s (\tilde{\Phi}_s^T \tilde{\Phi}_s)^{-1} \tilde{\Phi}_s^T \quad (3.10)$$

and α is the angle between the target subspace and the test vector and the matrix $\tilde{\Phi}$ contains the available a priori variability information about the target.

3.3.3 Parametric Classifier

Given a multivariate mixture model, with L classes of $\mathbf{w}_i, i=1, \dots, L$, the probability of pixel, at location \mathbf{x} belongs to class \mathbf{w}_i is given by the observation-conditional probabilities, $p(w_i | \mathbf{x})$, i.e the probability of class w_i given by the observation of pixel at location \mathbf{x} . Classification is performed by finding the class with maximum conditional probability (Kam, 2009):

$$\mathbf{x} \in w_i, \quad p(w_i | \mathbf{x}) > p(w_j | \mathbf{x}) \forall j \neq i \quad (3.11)$$

However, in practice these observation-conditional probability functions are often unknown.

If we have a sufficient number of training data, x_1, \dots, x_n for an accurate estimation, the probability of finding \mathbf{x} for each class can be assessed by estimating the probability distributions in each class using the Bayes's theorem (Kam, 2009),

$$p(w_i | \mathbf{x}) = p(\mathbf{x} | w_i)p(w_i) / p(\mathbf{x}) \quad (3.12)$$

Where $p(\mathbf{x})$ is the conditional probability that a pixel belongs to certain class and it is given by:

$$p(\mathbf{x}) = \sum_{i=1}^L p(\mathbf{x} | w_i)p(w_i) \quad (3.13)$$

From equation (3.12), $p(w_i | \mathbf{x})$ is known as the posterior probability and $p(w_i)$ is known as the prior probability. The prior probability for each class occurring is ($0 < p(w_i) < 1$) and for $i = 1, \dots, L$, the total prior probability is $\sum_{i=1}^L p(w_i) = 1$. Now, the classification rule of equation (3.12) is given by:

$$\mathbf{x} \in w_i, \quad p(\mathbf{x} | w_i)p(w_i) > p(\mathbf{x} | w_j)p(w_j) \forall j \neq i \quad (3.14)$$

with the common factor $p(\mathbf{x})$ removed (Kam, 2009). Since the logarithm is monotonically increasing the probability terms can be rewritten into:

$$g_i(\mathbf{x}) = \ln\{p(\mathbf{x} | w_i)p(w_i)\} = \ln p(\mathbf{x} | w_i) + \ln p(w_i) \quad (3.15)$$

where $g_i(\mathbf{x})$ is sometimes known as the discriminant function and the classification rules of equation (3.12) becomes

$$\mathbf{x} \in w_i, \quad g_i(\mathbf{x}) > g_j(\mathbf{x}) \forall j \neq i \quad (3.16)$$

3.3.3.1 Maximum likelihood classifier (QD)

In the case of Gaussian density with N spectral bands, the parameter for each class θ_i denotes mean m_i and covariance matrix Γ_i , $\theta_i = (m_i, \Gamma_i)$. The likelihood probability is defined by:

$$p(\mathbf{x} | w_i) = \frac{1}{(2\pi)^{N/2} |\Gamma_i|^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}_i - m_i)^T \Gamma_i^{-1}(\mathbf{x}_i - m_i)\right\} \quad (3.17)$$

The logarithmic form of the discriminant function becomes:

$$g_i(\mathbf{x}) = \ln p(w_i) - \frac{N}{2} \ln(2\pi) - \frac{1}{2} |\Gamma_i| - \frac{1}{2} (\mathbf{x}_i - m_i)^T \Gamma_i^{-1} (\mathbf{x}_i - m_i) \quad (3.18)$$

Since the $-\frac{N}{2} \ln(2\pi)$ term is constant for all $g_i(\mathbf{x})$, it can be removed to simplify the calculation. Often, there is no useful information about the prior probability and equal prior probability is assumed. By removing all the unnecessary constant terms, the final discriminant function can be refined:

$$g_i'(\mathbf{x}) = |\Gamma_i| + (\mathbf{x}_i - m_i)^T \Gamma_i^{-1} (\mathbf{x}_i - m_i) \quad (3.19)$$

Where

$$\mathbf{x} \in w_i, \quad g_i'(\mathbf{x}) < g_j'(\mathbf{x}) \forall j \neq i \quad (3.20)$$

This is sometimes known as the maximum likelihood classifier, log-likelihood classifier or quadratic (Gaussian) classifier (Kam, 2009; Dempster, *et al.* 1977).

3.3.3.2 Mahalanobis Distance classifier (MD)

If we assume that the covariance's Γ_i for all classes are equal that is $\Gamma_i = \Gamma$ for all class, the determinant of the covariance is constant and can be ignored. The discriminant function becomes

$$fd_i(\mathbf{x}) = (\mathbf{x}_i - m_i)^T \Gamma^{-1} (\mathbf{x}_i - m_i) \quad (3.21)$$

This is known as the Mahalanobis Distance classifier or Fisher Linear Discriminant classifier. A pattern is classified by finding the minimum distance from the normalised mean (Kam, 2009; Theodoridis, *et al.* 2010).

3.3.3.3 Euclidean distance classifier (ED)

Consider the covariance matrices of all classes to be diagonal and equal, and the variances in each component to be identical, therefore $\Gamma_i = \sigma^2 I$. The logarithmic form of the original log-likelihood discriminant function becomes

$$g_i(\mathbf{x}) = \ln p(w_i) - \frac{N}{2} \ln(2\pi) - \frac{1}{2} \sigma^{2N} - \frac{1}{2} (\mathbf{x}_i - m_i)^T \sigma^{-2} (\mathbf{x}_i - m_i) \quad (3.22)$$

Again we assume the prior probabilities are equal and remove all the constant terms, the discriminant function becomes

$$g_i(\mathbf{x}) = -(\mathbf{x}_i - m_i)^T (\mathbf{x}_i - m_i) \quad (3.23)$$

Here, we are trying to find the minimum $g_i(\mathbf{x})$ which is called the Euclidean distance. Therefore this type of classifier is called the Euclidean distance classifier or minimum distance classifier (Kam, 2009).

3.3.4 Detector employed in this work: MD classifier

During the course of this work the Mahalanobis Discriminant (MD) classifier, ACE and QD have been employed as detectors. Figure 3-5 shows the Receiver Operating Characteristics (ROC) results for the detection of a target (T3) from camera view #1 (denoted as T3C1) using these 3 detectors. It is seen from Figure 3-5 that the detection performances amongst these three detectors vary rather significantly. The MD detector has been chosen as the detector throughout this work.

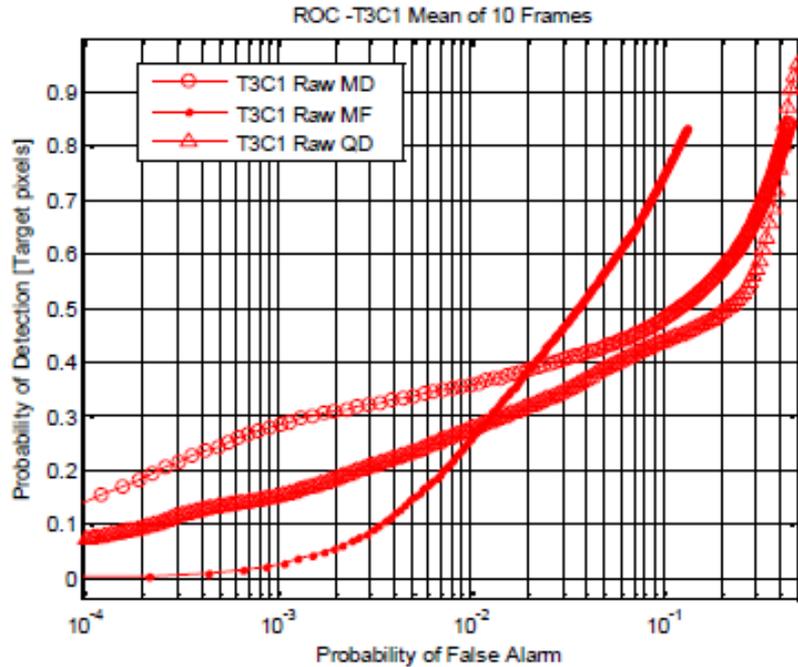


Figure 3-5: Shows the detection ROC for target T3 in camera view #1 (denoted as T3C1) using the 3 detectors MD (circle), QD (triangle) and ACE (dot) for the raw data. All detection was performed using the colour opponent descriptor.

3.4 Experimental Set Up

3.4.1 Target Signature Acquisition

The objective of this project is to examine how people detected in multiple non-overlap CCTV system could be improved by using colour constancy and spectral feature. Thus, there is no target segmentation or background subtraction involved in this work. All detections was performed using three dataset (as described above) namely MCT-TR (1001-1005) serials of the iLIDS (UK Home Office, 2008), PETS (2007) and GTCC datasets in this research work. The experimental procedure for target detection is carried out in the following procedures:

1. Short video sequence clips in which the chosen target makes appearance are pre-selected from all camera views as the test data clip.

2. A ROI box on the target is manually selected for the first THREE frames of the clips and the colour attributes inside the box is averaged as signature of each target. Normally the ROI box is selected from Camera view #1 and the signature is then stored in the memory for the detections of the same target in ALL camera views.
3. Clips of the video data are prepared in two forms: A) raw data B) the same clip but after processed by various CC algorithms.
4. Some targets do not have appearances in certain camera views and therefore detections will be performed on the views that they appear.
5. Detection is made for every 10 frames of video clips data and 10 detections (i.e. 10 frames of clip) are made for every target in each camera view. The averaged of these 10 detections (see Figure 3-6 Red plot) represents the overall detection performance of the target. The detection result is presented in ROC plotted for the probability of detection (PD) versus the probability of false alarm (PFA).

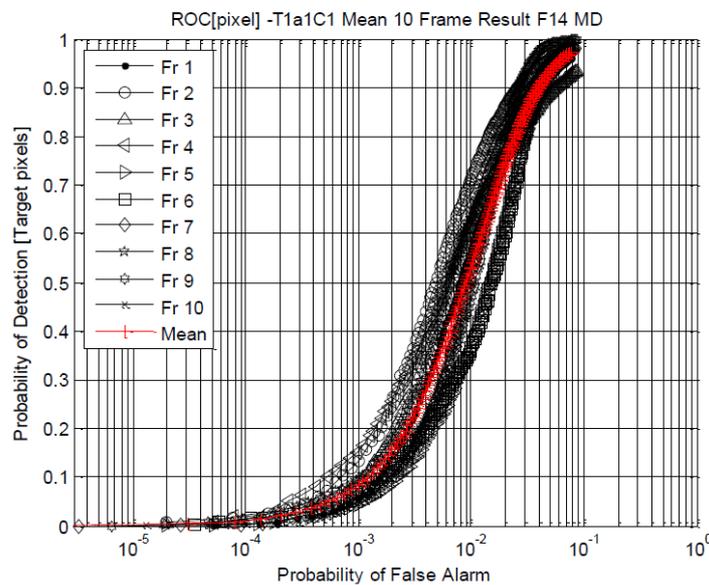


Figure 3-6: Typical ROC results for the detection of a target in every 10 frames interval (Black plots) of a video clip when the target transverse across the scene. There are 10 detections for a particular target and each produces one ROC. (Red plot) for better visualise the detection performance for particular target within video clip, all ROC results as seen in (Black plots) is averaged into one (Red plot).

3.5 Assessment of detection performances

Three methods for assessing the goodness of detection have been employed:

- 1) Figure of Merit (FOM) via ROC for target detection,
- 2) Delta Eab (ΔE_{ab}) and
- 3) Colour Difference Percentage (CDP).

3.5.1 Figure of merit (FOM) using AUROC

To quantify the goodness of detection, the area under the receiver operating characteristics (AUROC) between two fixed PFA values as shown in Figure 3-7, have been employed. The ROC assessment is widely used in area of target detection (Eisenbach, *et al.* 2012; Zhu, *et al.*, 2013) as a verification system. The AUROC is a more faithful way for assessing the detector performance rather than to compare the PD at a certain PFA value (Soori, *et al.* 2011). It is because most ROC do not shift rigidly as it is depicted in Figure 3-7 that at PD=0.4 the blue circled plot depicts ~2 order of magnitude better in PFA than that of the red curve, but it is only a factor of about 3 better at PD=0.7. The FOM is proposed in this research work (Soori, *et al.* 2013) and can be defined as equation (3.24):

$$FOM = \left(\frac{AUROC_{cc} - AUROC_{raw}}{AUROC_{raw}} \right) 100 \quad (3.24)$$

where the AUROC(raw) and AUROC(cc) are the area under the ROC for the raw untreated image data and that after it is processed by CC treatment respectively. Note that the all detection results, i.e., the ROC(raw) and ROC(cc), are obtained from the detections by using exactly the same detector (i.e. the MD classifier). The FOM is the performance index (in percentage) and the higher of the FOM the better of target detection enhancement after the CC treatment.

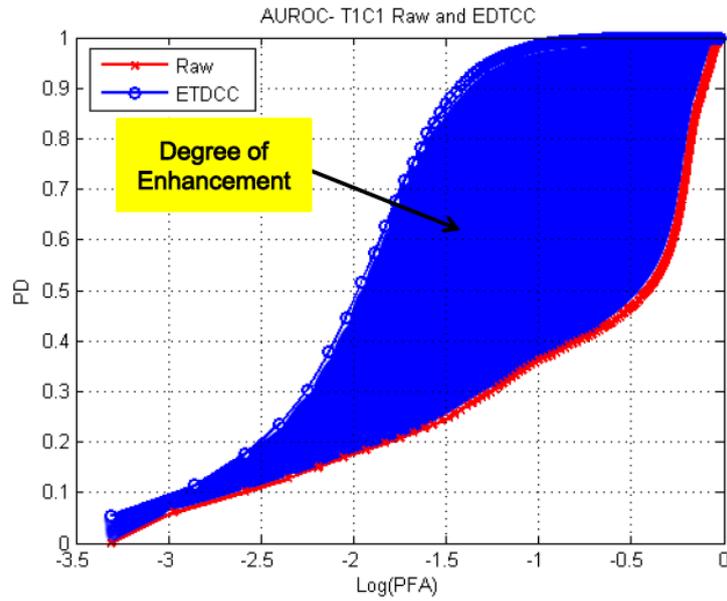


Figure 3-7: Illustrates how the enhancement of target detection is assessed in this work: the figure of merit (FOM) for the detection enhancement is evaluated by the difference in areas of the two ROCs with respect to the raw data according to equation (3.24). The area is calculated within the fixed PFA boundaries: in this case the enhancement is indicated by the area coloured in blue. Red trace: ROC of raw data, blue circle trace: ROC of the same data but after processed by ETDCC. The detectors in both cases are the Mahalanobis Discriminant (MD) classifier.

3.5.2 Delta ΔE_{ab}

Delta E_{ab} (ΔE_{ab}) has been widely used for assessing the colour correspondences between two given sample images. Euclidean Distance in L^*a^*b colour space has been commonly used for the ΔE_{ab} assessment (Sharma, *et al.* 2005):

$$\Delta E_{ab} = \sum_{n=1}^{n=10} \sqrt{(L_n - L_{10th(max)})^2 + (a_n - a_{10th(max)})^2 + (b_n - b_{10th(max)})^2} \quad (3.25)$$

where ΔE_{ab} is colour difference, L a b are the three bands of the image in L*a*b colour space, the subscript “n” denotes the test data set, subscript “10th(max)” denote the reference image. The ΔE_{ab} assessment has been applied for the GTCC dataset under the following conditions:

1. The 10th image which was acquired under uniform illumination at ~600 lux of irradiance is selected as the reference.
2. All images are converted into L*a*b and targets from 24 colour boxes are selected and their spectral characteristics are computed.
3. Test data: all images acquired under reduced irradiance (i.e. <~600 lux illumination).
4. ΔE_{ab} is evaluated.

3.5.3 Colour Difference Percentage (CDP)

CDP is a generalisation of ΔE_{ab} for other colour spaces:

$$CDP = \sum_{n=1}^{n=10} \sqrt{\left(\frac{B1_n - B1_{10th(max)}}{B1_{10th(max)}}\right)^2 + \left(\frac{B2_n - B2_{10th(max)}}{B2_{10th(max)}}\right)^2 + \left(\frac{B3_n - B3_{10th(max)}}{B3_{10th(max)}}\right)^2} \quad (3.26)$$

Where B_i is the value of band i.

Chapter 4: Enhanced Luminance Reflectance Colour Constancy Algorithm (ELRCC) For CCTV Network

This chapter details two modifications of luminance-based CC algorithms developed during the course of this research: 1. Colour transfer mechanism and 2. A pixel-wise sigmoid function for an adaptive dynamic range compression and it is also known as enhanced luminance reflectance CC (ELRCC). It has been found that both algorithms improve the efficiency of target detection substantially better than that of the raw data without CC treatment, and in some cases the ELRCC improves target detection by over 100% within the FOM assessment metric.

Most of the work presented in this chapter has been published in the OE paper (Soori, *et al.* 2013).

4.1 Colour Transfer (CT)

Camera calibration has been an important factor for maintaining the white balance and colour matching of the recorded images in each camera view within the multi-camera surveillance system. When this method is not available or when calibrations cannot be performed to achieve high degree of accuracy, then alternative method has to be employed for maintaining consistent colour characteristics over the entire multi-camera network. Colour Transfer (CT) is the method for the correction of colour differences in two sets of images using statistical mean and standard deviation of one of the camera data. This information is then transferred into other image scenes monitored by another camera. This method can be useful for applications such as in the image analysis of multi-camera system (Soori, *et al.* 2011). In real situations, however there are other factors such as variable illumination conditions, viewing angles and camera calibration settings which can induce colour distortions not intrinsic to the scene. Reinhard (2001) has developed a CT method for reducing illumination induced artefacts. The method utilised simple statistics of two images and introduced a relationship between the colour of the target image and that of the source image through a transform as shown in equation (4.1).

$$I'_{j(x,y)} = \mu_s + \left(\frac{\sigma_s}{\sigma_t} \right) (I_{j(x,y)} - \mu_t) \quad (4.1)$$

Where I' is the intensity map of the scene which receives the characteristics from another camera I (designated as source) within the network, j represents the (R, G, B) bands, s and t depict the information to be extracted from the source and target scenes respectively, σ and μ are the standard deviation and mean respectively. This method is termed as CT hereafter in this paper.

4.2 Enhanced Luminance Reflectance Colour Constancy Algorithm (ELRCC)

The proposed ELRCC method evaluates the luminance $L(x, y)$ at pixel (x, y) through the one band intensity $I(x, y)$ in the RGB colour space:

$$I(x, y) = \max(r(x,y), g(x,y), b(x,y)) \quad (4.2)$$

where r , g and b are the RGB components of colour images in RGB colour space. The intensity map I is the function of the luminance L and its reflectance R of the image scene as shown in equation (4.3):

$$I(x, y) = L(x, y) R(x, y) \quad (4.3)$$

There are several ways to estimate the luminance L of the scene and one approach uses a low-pass filtering of the intensities I at (m, n) through a 2D discrete Gaussian G as shown in equation (4.4)

$$L(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n) G(m+x, n+y) \quad (4.4)$$

where G is the 2D Gaussian at pixel location (x, y) and define as:

$$G(x, y) = q \cdot e^{-\left(\frac{x^2 + y^2}{c^2} \right)} \quad (4.5)$$

C is the surround neighbourhood constant ($c=2\sim 5$) and q is the normalisation constant and computed as:

$$\iint q.e^{\left(-\frac{x^2+y^2}{c^2}\right)} dx dy = 1 \quad (4.6)$$

Thus the reflectance R can be obtained through equation (4.3). Note that the L is not the real luminance as defined in physics but an approximated luminance that contains both the luminance and the mid-tone low-frequency components of the reflectance. Figure 4-1(b) shows intensity image of Figure 4-1(a) obtain by equation (4.2) and Figure 4-1(c) illustrates luminance image calculated by obtain by equation (4.4)

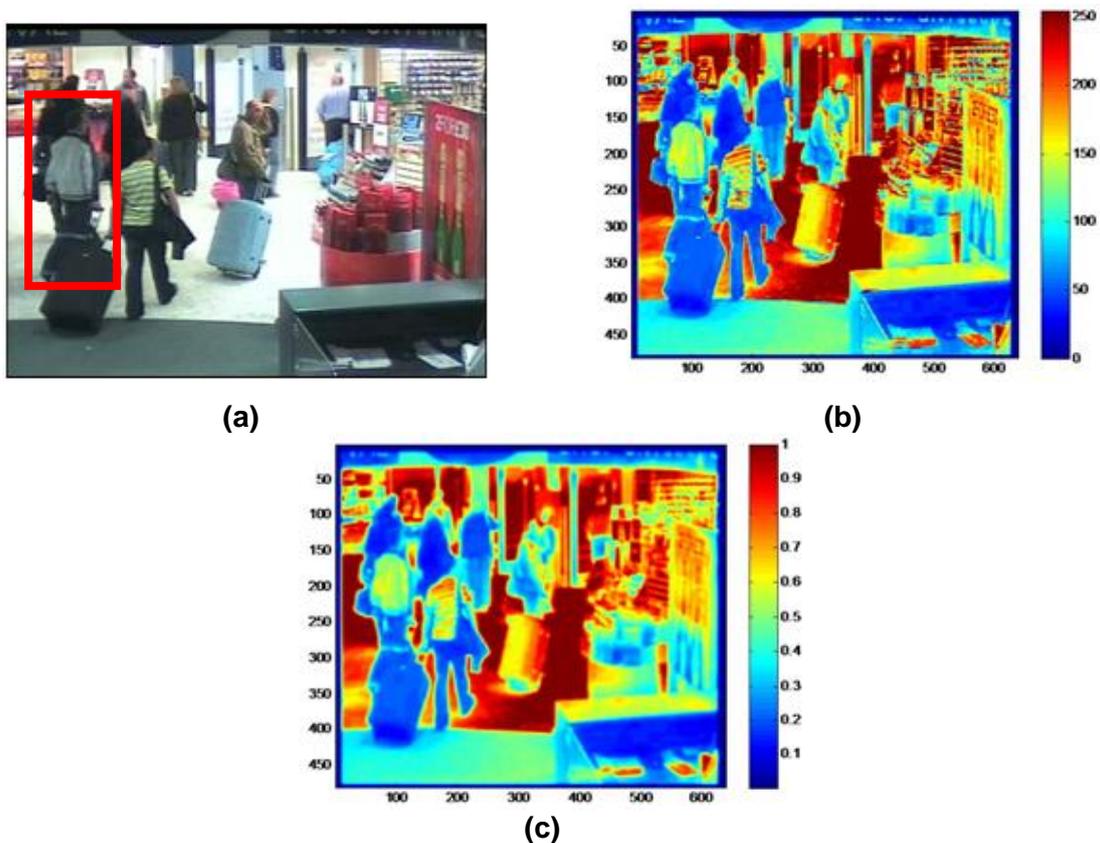


Figure 4-1: Highlights one band intensity and luminance images. (a) Original/Raw image of Target (red box) in Cam 1. (b) Intensity image obtain by equation (4.2). (c) Luminance image of (a) obtain by equation (4.4)

4.2.1 Adaptive Dynamic Range Compression

To achieve a balance of contrast enhancement and colour constancy, researchers in the field commonly use a non-linear transfer function, such as

the Windowed Inverse Sigmoid (WIS) function (Tao, *et al.* 2005), to condition the luminance such that the dynamic range of the image is compressed into a user defined moderate range. Typical WIS transfer function is shown in equation (4.7)

$$f(v) = \frac{1}{1 + e^{-v}} \quad (4.7)$$

In practise the luminance is moderated into a range of user defined v_{\min} and v_{\max} through the transfer function:

$$L_n' = L_n [f(v_{\max}) - f(v_{\min})] + f(v_{\min}) \quad (4.8)$$

$$L_n'' = \ln\left(\frac{1}{L_n'} - 1\right) \quad (4.9)$$

$$L_{n,enh} = \frac{L_n'' - v_{\min}}{v_{\max} - v_{\min}} \quad (4.10)$$

where equation (4.8) is employed for linearly mapping of input image to the magnitude range $f(v_{\min})$ $f(v_{\max})$ for equation (4.9), which is the non-linear inverse sigmoid function. Equation (4.10) is used to normalise the intensity transfer to bind the range of the luminance. Parameters v_{\max} and v_{\min} controls the shape of the transfer function. The range of v_{\min} and v_{\max} effect the contrast enhancement, and the v_{\max} can be fixed as arbitrary, while v_{\min} can be extracted from the scene. The method described in (Tao, *et al.* 2005) has been using global mean and in this work it is proposed to evaluate the v_{\min} for every pixel. This will give more control over the details of colour correction: the dark pixels should be given a smaller v_{\min} value whereas bright pixels should have larger value of v_{\min} . To achieve this objective equation (4.11) is implemented:

$$v_{\min} = \left(\left(\frac{I(x, y) - 1}{254} \right) k \right) - 6 \quad (4.11)$$

where k is a user define fixed constant and from experiments the appropriate value of v_{\min} is selected from -1 to -6 and $k = 5$ satisfying the equation (4.11) for v_{\min} -1 to -6 for 8-bit images, thus $k = 5$ has been used throughout this work.

4.3 Effects of V_{\min} in Luminance Based CC Algorithm

The v_{\min} effects the dynamic range and thereby it has direct influence on the brightness and contrast of the image. Previous work (Tao, *et al.* 2005) set this parameter for the whole image as a user's define constant, which can be obtained from the global mean of the scene. Figure 4-2 (b) & (c) are the results of one clip of the raw video sequence as shown in Figure 4-2 (a), after processed by the band wise luminance based CC using the $v_{\min} = -5$ and -10 respectively. The RGB attributes of the target within the ROI which is depicted as the red rectangle in the images (a-c), are shown in Figure 4-2 (d) and (e) after the video sequence is processed using constant v_{\min} of -5 and -10 respectively. The sampling frequency is 5 frames per sec, and the colour attributes are in RGB. It is quite clear that the colour constancy is rather poor when $v_{\min} = -5$ (Figure 4-2(d)), and it improves a little when $v_{\min} = -10$ is used (Figure 4-2(e)). However, the larger value of the v_{\min} also increases colour bleaching.

Intuitively, a pixel wise v_{\min} is needed for the colour constancy due to the non-uniform illumination across the scene. Figure 4-3 compares the effects of using a constant v_{\min} for the global scene (Tao, *et al.* 2005) with that using a pixel wise implementation according to equation (4.11). The colour characteristics of the target after processed by the pixel wise v_{\min} are shown in Figure 4-3 (d), indicative of having a much better colour constancy performance than that using a global v_{\min} for the whole scene (Figure 4-3 (c)).

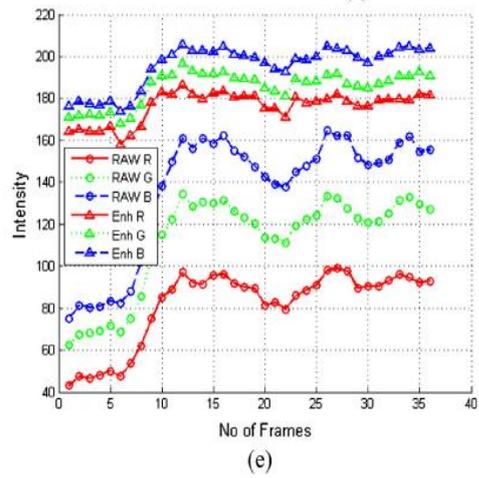
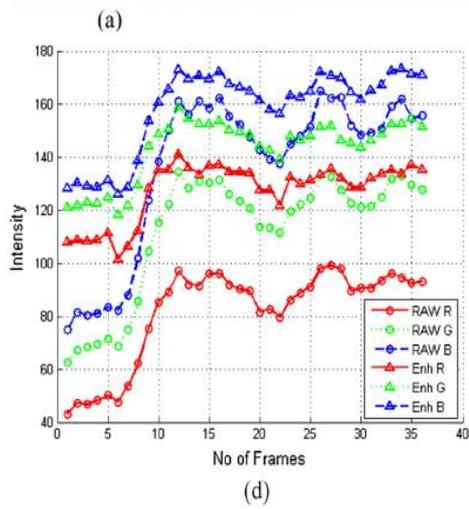


Figure 4-2: Illustrates the effects of v_{min} settings on the colour constancy of a video sequence after processed by the luminance based algorithm. (a) Raw data, (b) $v_{min} = -5$ and (c) $v_{min} = -10$, (d) colour attributes of target (in red rectangle) after processing data using $v_{min} = -5$ and (e) colour attributes of target after processing data using $v_{min} = -10$.

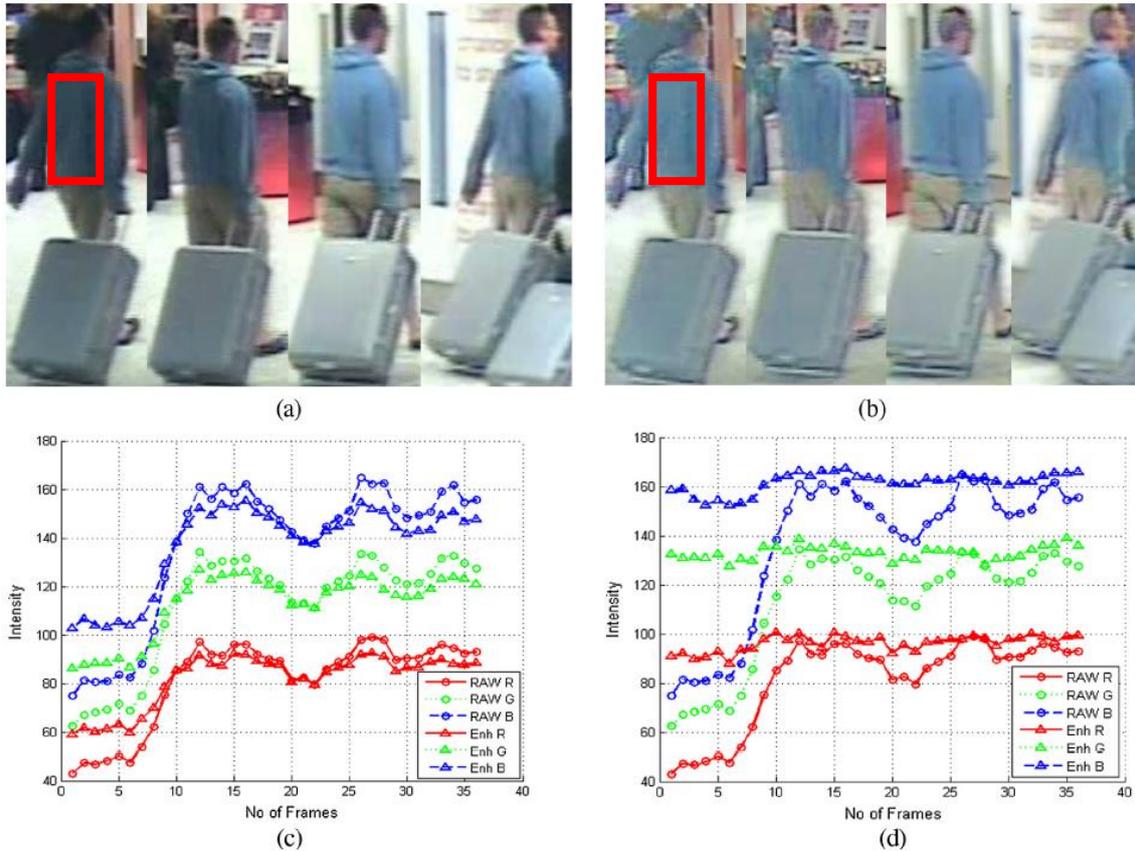


Figure 4-3: Shows the effect of pixel wise dynamic range compression in the luminance based CC algorithm: (a) & (c) after processed by using a global mean v_{min} for the complete scene as proposed by (Tao, *et al.* 2005), (b) & (d) after processed by the proposed pixel wise v_{min} . (c) & (d) depict the colour characteristic of the target. Note that the images that have been processed by the global v_{min} in (a) is found not capable to rectify local colour non-uniformity due to variable illumination conditions.

4.3.1 Adaptive Mid Tone Frequency Components Enhancement

Like all the Retinex algorithms, the luminance as depicted above exhibits a mid-tone and low frequency components which can be degraded by the dynamic range compression. A centre-surround type of contrast enhancement method can be utilised to help compensate this degradation as shown in equation (4.12):

$$L'_{n,enh}(x, y) = L_{n,enh}(x, y)^{E(x, y)} \quad (4.12)$$

where the exponent $E(x, y)$ is defined by:

$$E(x, y) = \left(\frac{I_{conv}(x, y)}{I(x, y)} \right)^P \quad (4.13)$$

where $I(x, y)$ and $I_{conv}(x, y)$ are the luminance evaluated using larger values of the neighbourhood constant c (10~20) in equations (4.4 - 4.5). The $L'_{n,enh}(x, y)$ is the enhanced luminance image after pixel-wise sigmoid function. The exponent P is chosen to be a function of the global standard deviation in the image $I(x, y)$ which measures the extremeness of the intensity map. The exact value of P is scene dependent and it can be determined by experiments. Typical values of P is in the range of (0.1~1) for the data set utilised in this study. Given the knowledge of the luminance L , then the reflectance R can be achieved using the relationship from equation (4.3). Figure 4-4(b) shows reflectance image. Once the final luminance $L'_{n,enh}(x, y)$ is obtained the enhanced one-band image $I_{enh}(x, y)$ can be obtained by combining the reflectance R from equation (4.3) in equation (4.14). Figure 4-4(c) shows the one-band enhanced intensity $I_{enh}(x, y)$.

$$I_{enh}(x, y) = L'_{n,enh}(x, y) R(x, y) \quad (4.14)$$

The process for arriving the $I_{enh}(x, y)$ is termed as ELRCC (Enhanced Luminance Reflectance Colour Constancy algorithm) hereafter in this thesis.



Figure 4-4: Outlines the ‘reflectance’ and one-band intensity map: (a) the intensity $I(x,y)$ of original image for Target (red box) in Cam1 (b) luminance based ‘reflectance’ image obtained by equation (4.3) (c) the one-band enhanced intensity I_{enh} by using equation (4.14).

The colour information is stored in a single channel image $I_{enh}(x, y)$ via equation (4.15).

$$\begin{aligned}
 ELRCC_{(R)}(x, y) &= (I_{enh}(x, y) / I(x, y)) R(x, y) \\
 ELRCC_{(G)}(x, y) &= (I_{enh}(x, y) / I(x, y)) G(x, y) \\
 ELRCC_{(B)}(x, y) &= (I_{enh}(x, y) / I(x, y)) B(x, y)
 \end{aligned} \tag{4.15}$$

where $ELRCC_{R,G,B}(x, y)$ is the result of the 3 colour channels after processed by the ELRCC algorithm in the RGB colour space.

4.4 Results

4.4.1 Retinex and CT Results

Figure 4-5(a) represents the raw image with target (red box) presents in the Camera 1 view. Figure 4-5 (b), (c) and (d) shows result image after application of SSR where Gaussian surround function “c” equal to 15, 80 and 250 respectively. It is found that SSR is incapable of achieving CC regardless of using different surround function “c”.

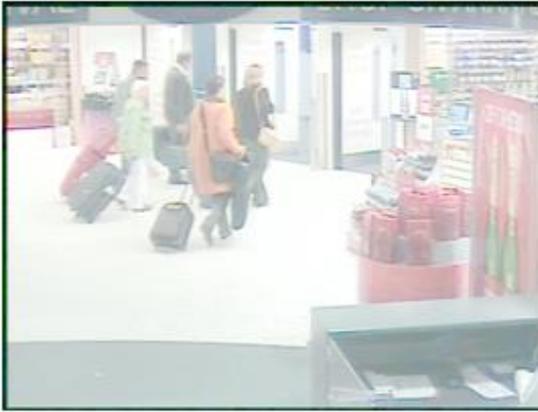
The same data but processed by MSR algorithm is shown in Figure 4-6(b) which have utilised the Gaussian surround constants of $c_1=15$, $c_2=80$ and $c_3=250$. The results are seen to improve in comparison to that of the SSR, but the colour integrity of the processed image is degraded too. Note that this MSR can be generated by linear combination of the SSR results as shown in equation (2.1).

The same data is applied using MSRCR as shown in Figure 4-7(b) using the same surround constants of $c_1=15$, $c_2=80$, $c_3=250$. It is seen that the colour correction improves the colour compensation when it is compared with Figure 4-6(b) (MSR without colour correction). The gain and offset are not very sensitive to the colour restoration, but the alpha and beta in equation (2.6) can change the colour attributes rather significantly (Figure 4-7).



(a)

(b)



(c)



(d)

Figure 4-5: Shows SSR result. (a) Raw/original Image with Target (red box) in cam 1. (b) SSR result with Gaussian surround constant “c” = 15. (c) SSR result with Gaussian surround constant “c” =80. (d) SSR result with Gaussian surround constant “c” = 250.



(a)



(b)

Figure 4-6: Shows MSR result. (a) Raw/original Image with Target (red box) in cam 1. (b) MSR result with commonly used Gaussian surround constants $c_1=15$, $c_2=80$ and $c_3=250$.



Figure 4-7: Shows MSR-CR result. (a) Raw/original Image with Target (red box) in cam 1. (b) MSR-CR result with surround constants $c_1=15$, $c_2=80$, $c_3=250$, gain $G=192$, offset $b=-30$, $\alpha=125$ and $\beta=46$ (c) MSR-CR result with surround constants $c_1=15$, $c_2=80$, $c_3=250$, gain $G=50$, offset $b=-3$, $\alpha=125$ and $\beta=46$ (d) MSR-CR result with surround constants $c_1=15$, $c_2=80$, $c_3=250$, gain $G=192$, offset $b=-30$, $\alpha=10$ and $\beta=5$

The performance of Sub-Band Decomposition MSR (SBD-MSR) algorithm is shown in Figure 4-8 which presents the raw and the processed data using SBD-MSR algorithm on the left and right hand panels respectively. In (a), (b) & (c) it shows the image of a target (red box) extracted from Camera 1, 3 and 5 respectively. It is quite obvious that the colour of images in all three camera views have been washed out significantly and resulting in grey. Even for strong colour like the pink coat, the colour is seen to reduce to very pale making it very difficult to recognise after the SBD-MSR processing.



(a)



(b)



(c)

Figure 4-8: presents the raw and the processed data using SBD-MSR algorithm respectively in the left and right hand panels. In (a), (b) & (c) it shows the images of a target (red box) in clips extracted from Camera 1, 3 and 5 respectively.

The LB-MSR algorithm was coded in MATLAB according to the equations set out in section 2.1.2.2, and the luminance of the scene according to equation (2.12) is illustrated in Figure 4-9, which is plotted such that high luminance is represented by 'hot' colours. Note that in LBMSR algorithm the colour constancy is 'conditioned' according to the luminance of the scene and this is the main difference with the conventional MSR which relies heavily on the appropriateness of the user input parameters.

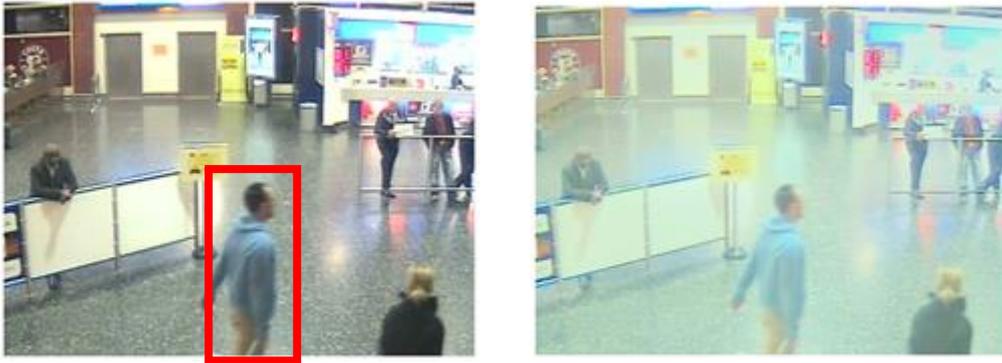


Figure 4-9: Illustrates the luminance (in right) of an image (left).

The colour attributes are seen to be somewhat less degraded in the LBMSR processed images in comparison to that of the MSR. Figure 4-10 shows the raw (left) and processed data by LB MSR (right) for a target (red box) that had made appearance in camera 1, 3 and 5. The distinct difference between the LB MSR (Figure 4-10) and the MSR (Figure 4-8) results is the more superior colour retention in the luminance based methodology.



(a)



(b)



(c)

Figure 4-10: shows the raw (left) and processed data by LB MSR (right) for a target (red box) who had the appearance in (a) camera 1, (b) camera 3 and (c) camera 5.

4.4.2 ELRCC Results

4.4.2.1 Targets Representations

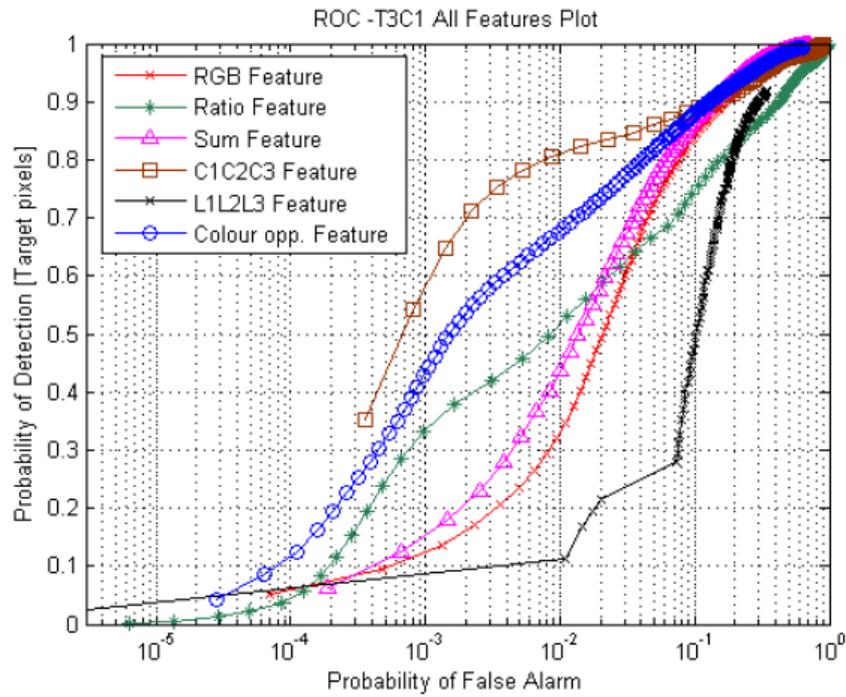
Ten subjects with various coloured clothing have been chosen as targets in this experiment and typical RGB images of them are shown in Figure 4-11.



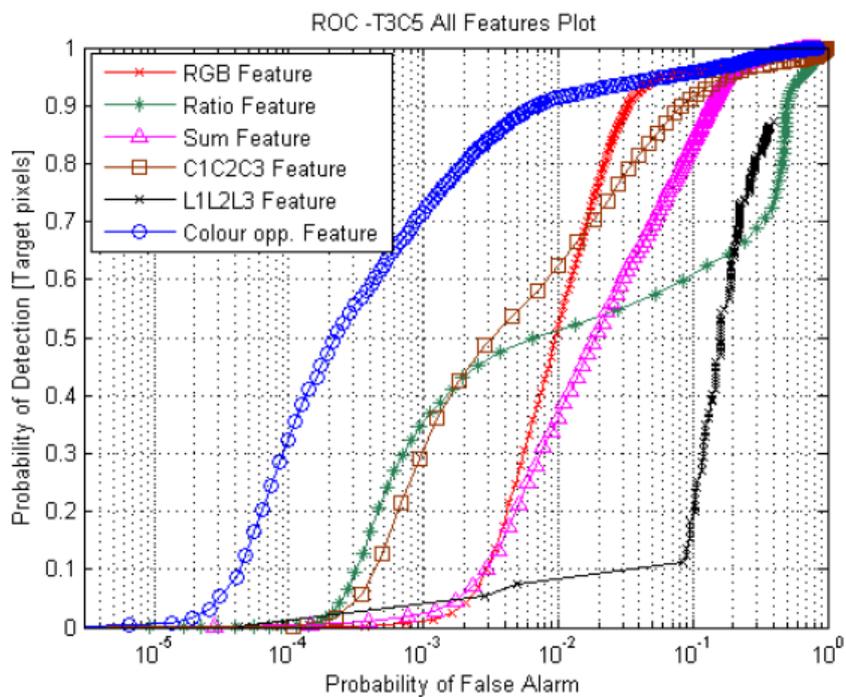
Figure 4-11: Representative RGB pictures of 10 selected targets in this chapter: (left to right) T1 to T10, respectively.

4.4.2.2 Optimum colour descriptor

It is well known that the selection of inappropriate colour feature descriptor is detrimental to object detection. To understand the effectiveness of the colour descriptor for maintaining colour constancy, we have employed 6 different colour descriptors in this study: RGB Feature, Ratio Feature (Soori, *et al.* 2011), Sum Feature (Soori, *et al.* 2011), L1L2L3 descriptor (Gevers & Smeulders, 1999), C1C2C3 descriptors (Gevers & Smeulders, 1999) and colour opponent feature (Gijsenij & Gevers, 2011). Figure 4-12 present the averaged results of 10 frames of images for the detection of target T3 in camera 1 view using these six different colour descriptors. It is seen that the colour opponent feature and the C1C2C3 descriptors have exhibited rather good colour invariance properties with better detection performances than that of all other descriptors. Figure 4-12 also highlights the importance of the descriptors: the good descriptors, for example, C1C2C3 & colour opponent feature, can reduce the PFA by two orders of magnitude in comparison to the bad one, such as the L1L2L3. It is also found that the C1C2C3 descriptor cannot give consistent colour invariance performance (see Figure 4-12 (b)) due to chrominance conversion (normalisation) the C1C2C3 descriptor unable to distinguish different shade of grey (see section 6.2 for more detail) and therefore the colour opponent has been selected for assessing the effectiveness of various colour constancy algorithms in this chapter.



(a)



(b)

Figure 4-12: Shows the averaged detection ROC of T3 ELRCC data in (a) camera view #1 (b) camera view #5. Note that some descriptors, such as the L1L2L3, have exhibited very poor colour invariance property thereby giving rather poor detection performances. The C1C2C3 descriptor also gives very variable results which can be seen from the ROC of the two camera views shown in (a) & (b).

4.4.2.3 Target detection performances of various CC algorithms

In this section, the effectiveness of various forms of CC algorithms is testified here through their object detection performances for using one video clip set as the test data. The video clip set is firstly processed by all these CC algorithms, and their Receiver Operating Characteristics (ROC) and subsequently their Figure of Merit (FOM) is compiled to give some indications whether they have improved target recognitions with respected to the untreated raw data. The experiment involves the detection of a randomly selected target, and in this case the target T3, who had made appearances in camera views #1, #3 & #5 . These video clips are also employed as the test data of this experiment. Six different CC algorithms have been employed:

- i) the proposed ELRCC (Soori, *et al.* 2013),
- ii) the luminance based and CT method (Soori, *et al.* 2011) (denoted by LB CT),
- iii) the MSRCR (Rahman, *et al.*1998),
- iv) the LBMSR (Tao & Asari, 2003),
- v) the luminance based method which uses global mean depended fixed v_{min} for the entire scene (Tao, *et al.* 2005) (denoted by Fixed V_{min});
- vi) the sub-band MSR method band wise (Jang, *et al.* 2008) (denoted by SubBand MSR).

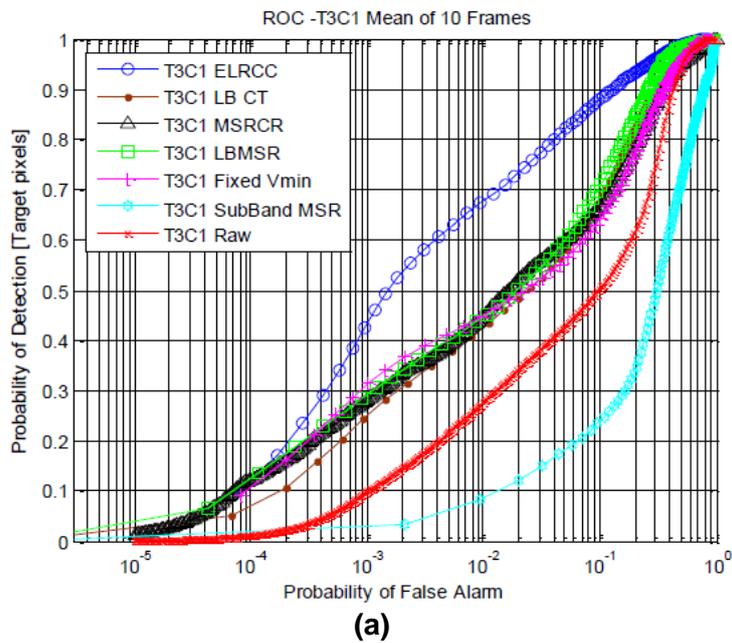
In all cases the colour opponent feature has been employed throughout and the detection was performed using the MD classifier. Figure 4-13 shows the detection results of target T3 for the 3 different camera views after the image data is processed by the six different CC algorithms as mentioned above.

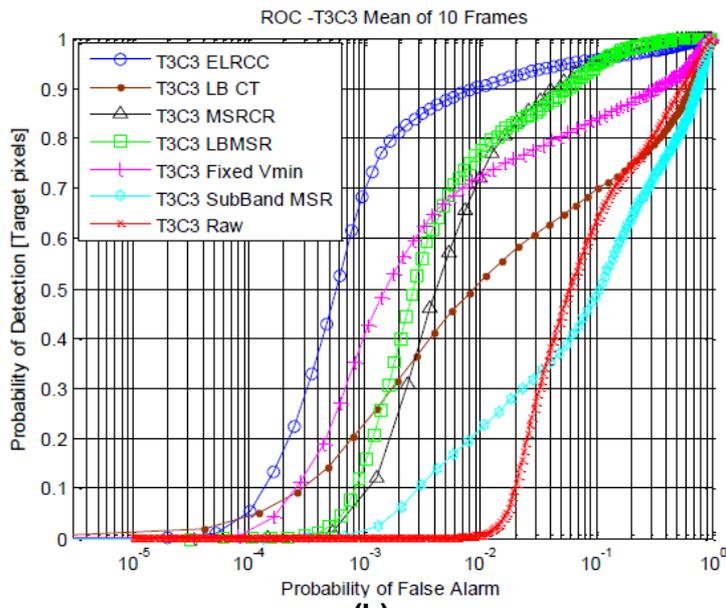
Two main results have been observed from Figure 4-13:

1. All CC algorithms apart from the proposed ELRCC are not producing consistent performance and in many cases the detections are not even as good as that of the raw data.

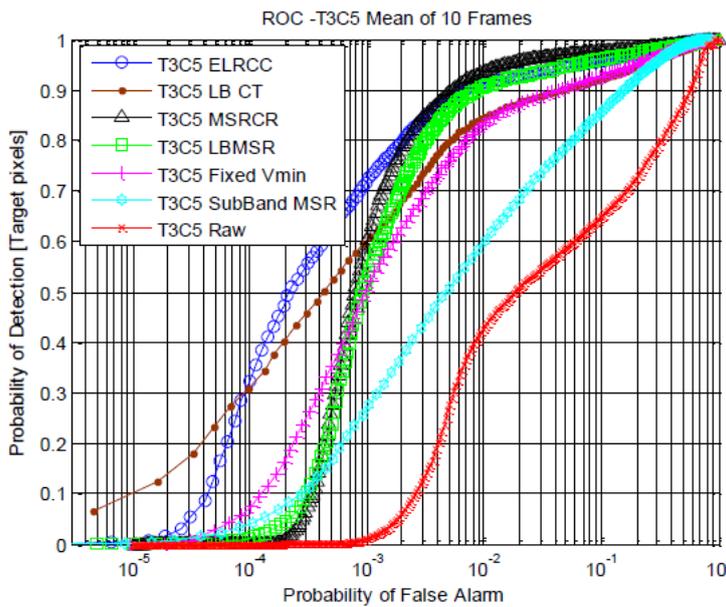
- The ELRCC algorithm shows better the target detection rate at low probability of false alarm (PFA). For example Figure 4-13(a) shows under 40-45% of probability of detection (PD) for all other CC methods at the PFA 10^{-2} , while the data processed by ELRCC shows ~70% of PD which exhibits ~25-30% higher detection rate than all other methods.

Table 4-1 tabulates the FOM for the detection of Target T3 in 3 different views after the data is treated by these six different CC algorithms. It is seen from the table that the CT method barely improves target detections, while the proposed ELRCC enhances target detections by about a factor of 2 better than all other CC methods utilised in this experiment.





(b)



(c)

Figure 4-13: Average of 10 detection results for target T3 in (a) camera view #1, (b) camera view #3, and (c) camera view #5, after the video sequence is processed by six different CC algorithms. Note that all CC algorithms, except for the one enhanced by the proposed ELRCC (in blue circle) method, exhibit little or no detection improvements with respect to that of the raw untreated data (in red cross) in some camera views.

Table 4-1: Tabulates the Figure of Merit (FOM) in % for the detection of Target T3 in 3 different views after the data is treated by six different CC algorithms.

	CC Methods	Camera 1	Camera 3	Camera 5	Mean FOM in %
1	ELRCC	91.2	178.52	126.28	132
2	LB CT	35.35	68.78	110.56	71.56
3	MSRCR	46.09	114.22	92.04	84.12
4	LBMSR	52.01	115.63	88.96	85.53
5	Fixed V_{min}	40.94	117.5	84.39	80.94
6	SubBand MSR	-49.21	3.61	48.17	0.86

4.4.2.4 ELRCC algorithm: 10 targets test

The results that have been presented in the last section may suggest a better colour constancy capability in the proposed ELRCC algorithm, than all other 5 different CC algorithms employed in this study. However, this is only based on one result from the detection of target T3. In this section, the capability of the ELRCC algorithm is critically assessed here by using data sets of 9 other targets. As previously, the detection is made for every 10 frames of video clips and the detection results for each target in each camera views are presented in ROC which is then averaged to form a representative ROC. 4-14 shows the typical RGB images of three camera views (view #1, view #3 and view #5) of target T10 with the raw data shown in the upper panel, the same data but after processed by ELRCC is shown in the lower panel. Note that the colour bleaching is very minimal after the data is processed by the ELRCC algorithm.



(a)



(b)

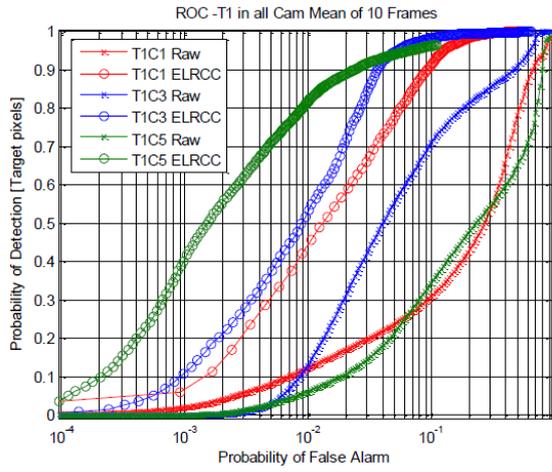
Figure 4-14: Shows representative RGB images selected from 3 camera views of target T10: (upper panel) raw data and (bottom panel) after processed by the ELRCC algorithm.

Due to the large number of targets utilised in this experiment, only the detection results for the first five targets (T1 to T5) are presented here, and their ROC results are depicted in Figure 4-15 (a) to (e) for these five targets respectively. The ROCs are colour coded, depicting the results from camera views #1, #3 and #5 in red, blue and green respectively. It is quite clear that the ELRCC processed data (in circle trace) exhibits consistently improved target detections over that of the raw data (in dot trace) for all camera views. The degree of enhancements can be indicated by the shift of the ROC curve to the right which can be measured more quantitatively from the area of the ROC within the PFA bounds of $[10^{-5}, 1]$. The enhancement of target detection after the data is processed by ELRCC is computed for all 10 targets for all camera views using the FOM formulation according to equation (3.24). Note here that the T5 in camera view #3 in Figure 4-15(e) shows quite the bad detection results as compare to other camera views, it is mainly due to high false alarm rate because cam #3 background have similar spectral signature as the target.

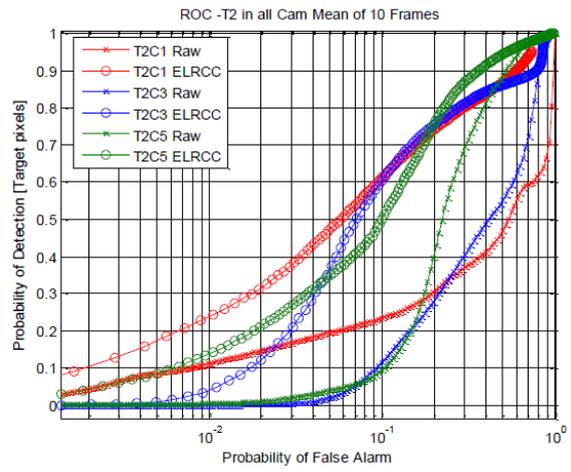
The result is tabulated in Table 4-2 which exhibits substantial detection enhancement for the data that has been processed by the proposed ELRCC algorithm. In many cases there are over 100% of detection enhancements that have been seen particularly from camera view 5. Note that all targets in all camera views have shown variable degrees of detection enhancements after the data is processed by the proposed ELRCC algorithm.

Table 4-2: Tabulates the FOM for the detection performance enhancement for the data processed by the proposed ELRCC for all 10 targets.

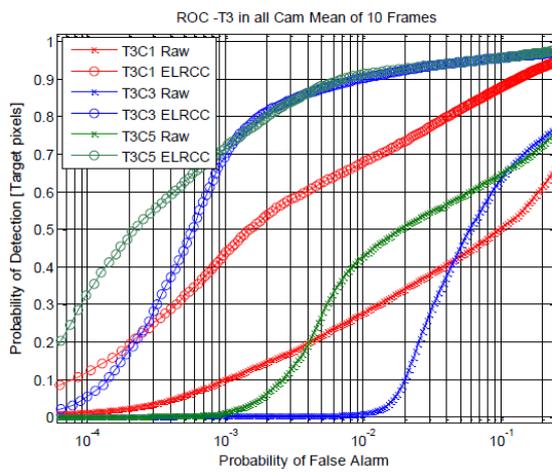
Targets	Camera 1	Camera 3	Camera 5	Mean FOM In %
T1	82.95	45.73	213.45	114.04
T2	108.98	123.25	82.59	104.94
T3	91.20	178.52	126.28	132.00
T4	62.38	109.51	129.04	100.31
T5	43.69	63.89	122.12	76.57
T6	37.17	50.55	78.66	55.46
T7	91.56	74.83	199.80	122.06
T8	43.52	Target Absent	89.64	66.58
T9	86.81	48.35	149.21	94.79
T10	69.29	Target Absent	108.51	88.90
Averaged	71.75	86.83	129.93	96.17



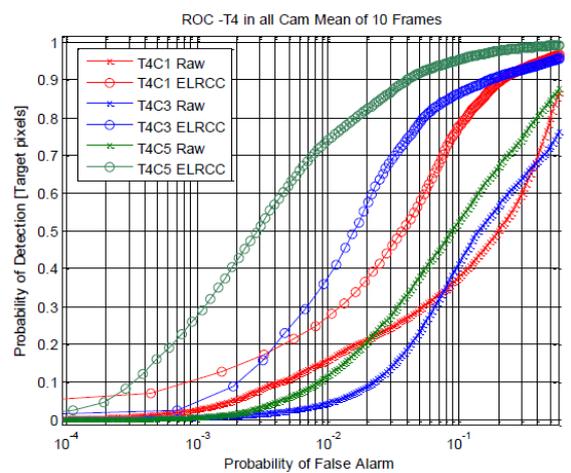
(a)



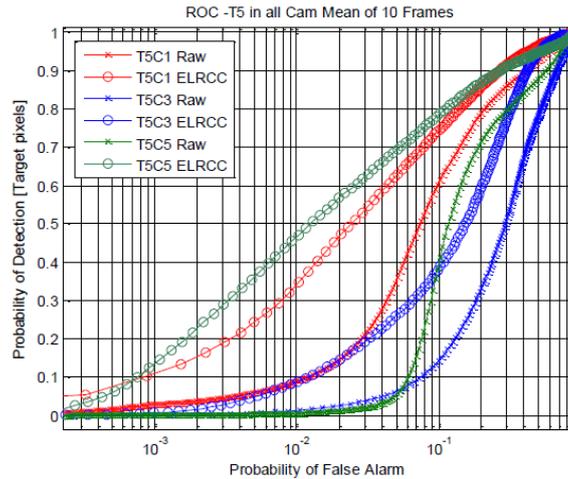
(b)



(c)



(d)



(e)

Figure 4-15: Shows the mean of the ROC results for the detection of targets (a) T1, (b) T2, (c) T3, (d) T4 and (e) T5. The ROC of the raw data and that after processed by ELRCC are presented in circle and cross respectively. The three different camera views of data are presented in colour code of red, blue and green for view #1, view #2 and view #3 respectively. All detection was performed using the colour opponent descriptor.

4.4.3 Discussion

The main result in this chapter is presented in Figure 4-15 and Table 4-2, which have shown a consistent trend of detection enhancements for all 10 targets after the data is processed by the proposed ELRCC algorithm. This result may suggest that the CC processing can indeed help to reduce illumination artefacts and to help restore the colour integrity of the scene. The averaged target detection enhancement over the 10 target in the 3 camera views #1, #3 and #5 are found to be ~71%, ~87% and ~130% respectively in FOM metric. However, more of the enhancement is seen from the detections in camera view #5, where the range of detection is much further amongst all the three views presented in this work. Furthermore, the rear view of the scene in camera #5 was greatly affected by the strong solar illumination through the window, and therefore the detection is expected to be more difficult particularly when the spectral characteristics of the very limited number of the target pixels are heavily distorted by the illumination artefacts. It is also noted that the least detection enhancement is seen from the data recorded by camera view #1, where the

target signatures were taken from the first 3 frames of the video sequence. The relatively poor detections may be due to the more extreme change of illumination condition inherit in this camera view, and partly may be due to the presence of many colourful objects in the background of the scene.

The average detection enhancement for all three camera views over all the 10 targets is found to be ~90% in FOM metric. In order to visualise the implication of this enhancement as the result of the CC treatment by the proposed ELRCC algorithm, all the ROCs that presented in Figure 4-15, together with that of the rest of the other 5 targets, are summed and averaged as shown in Figure 4-16. It is quite clear from this figure that the ROC of the CC processed data has shifted rather rigidly to the lower PFA rather substantially. At PD=0.5 the PFA reduction as the result of the CC process by the proposed ELRCC is found to be about a magnitude better than that of the raw untreated data. It is to emphasise that this amount of PFA reduction is the averaged result for over all ten targets in three different camera views.

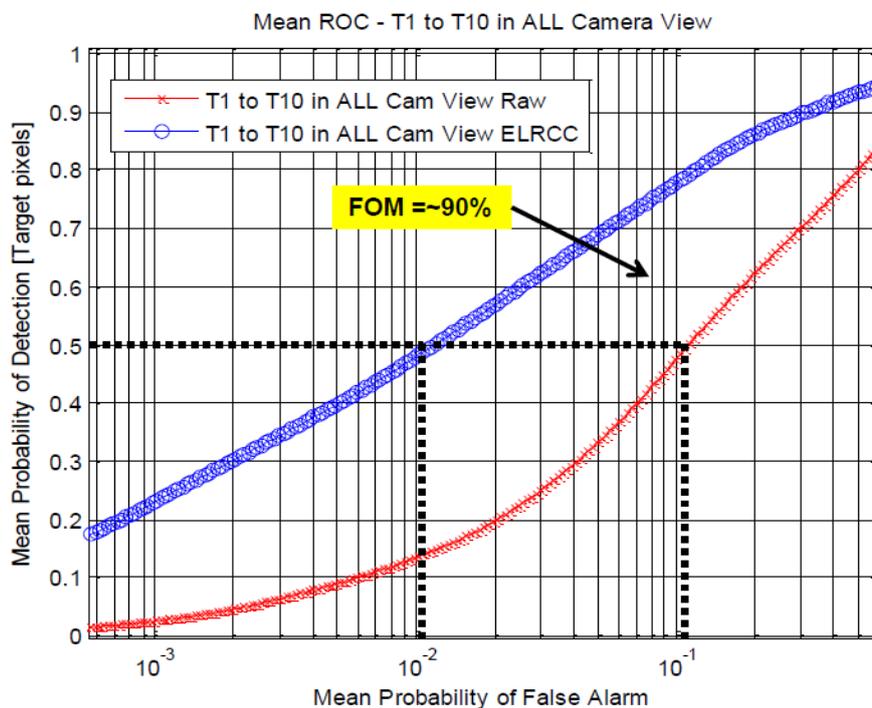


Figure 4-16: Shows the averaged ROC detection results for all 10 targets in all camera views: the averaged raw ROC in red trace, and the average ROC for the ELRCC treated data in blue which shows a FOM of target enhancement by 90% FOM, implying for a PFA reduction by about a magnitude at PD=0.5.

Chapter 5: Enhanced Target Detection Colour Constancy algorithm (ETDCC)

The main drawbacks of Retinex based method is the “user defined” parameterisation such as the spatial scale constant, Gaussian surround function parameters, α the strength of non-linearity, β the control constant, gain G and offsets b for colour restoration are needed to be tuned manually. Furthermore MSR methodology is seen to suffer from these issues: 1) Noise in the dark region which tends to alleviate further in the processing chain. 2) Artefacts at the boundary between the dark and bright areas. 3) Dark and shadow regions remains dark if the image have higher proportion of brighter regions. 4) Computational cost is quite high due to convolution between large images. There have been attempts for the parameterisation using edge sharpness, (Orsini, *et al.* 2003) and automatic tuning (Ciurea & Funt, 2004) but these techniques lacks in robustness and hence more research in this direction is needed.

The main cause of these issues is due to the use of global parameterisation in these algorithms. One solution to these problems is the pixel-wise parameterisation approach as detailed in the last chapter (Soori, *et al.* 2013).

Target detection becomes even more challenging and demanding in the CCTV system where many parameters/settings are unknown. It is necessary to take multiple measures in order to enhance target detection in this complex environment. Some illumination invariant colour features (such as rgb and C1C2C3 descriptors) are found applicable only for the targets with strong colours (for detail discussion section 6.2). Thus a multiple approach consisting of:

1. A bidirectional pixel-wise non-linear transfer functions for mid-tone compression. The transfer function also acts as luminance synchronisation across different video frames and across different camera views. This algorithm is one of the contributions from this work.
2. Luminance enhancements using centre-surround convolution.

3. Grey Edge algorithm is utilised mainly for colour correction and white balancing in the image scene.

Figure 5-1 show the flow chart of the proposed ETDCC algorithm.

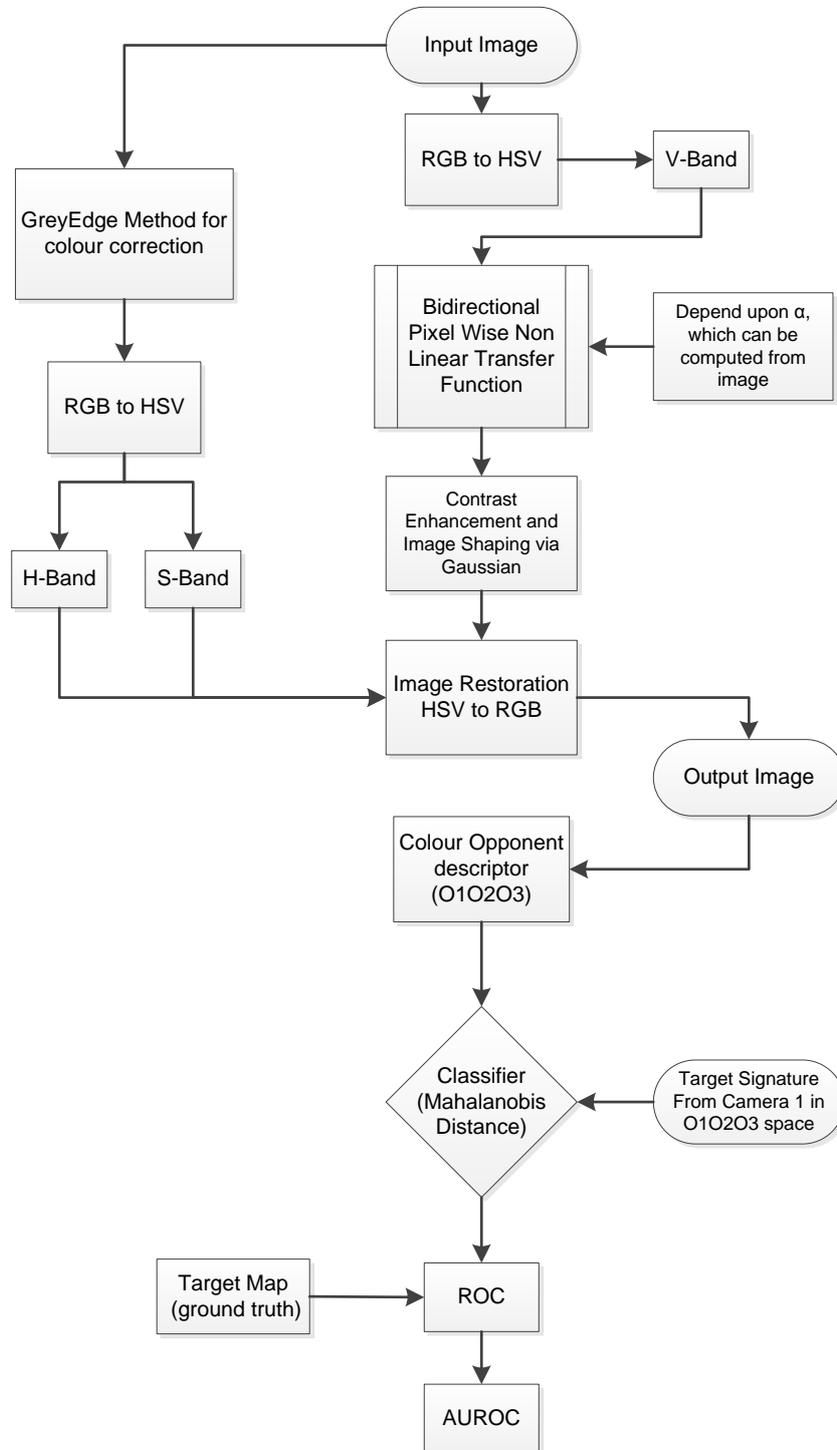


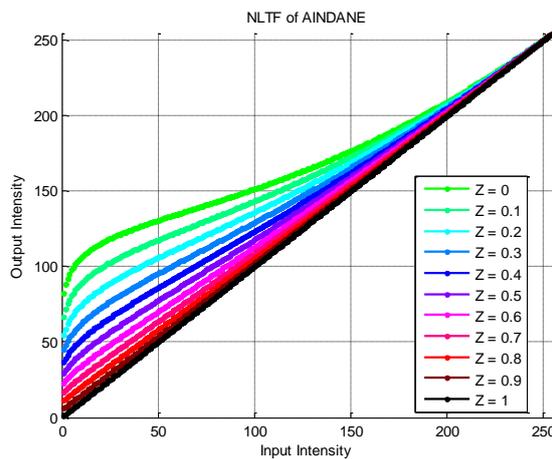
Figure 5-1: Outlines the flowchart of the proposed ETDCC algorithm.

5.1 Non-Linear Transfer Function (NLTF)

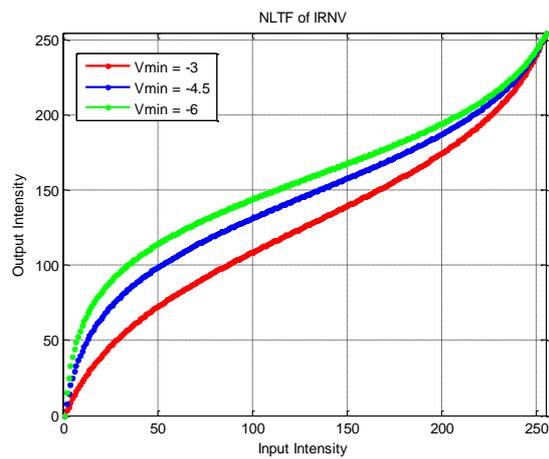
This is the first step in the work flow.

Dark pixels in the CCTV footage can be partially corrected using techniques such as gamma adjustment, histogram equalisation and the AINDANE (Tao & Asari, 2005) type of algorithms. However these types of methods utilised global parameter which do not have appreciable effects for small patches of dark/bright ROIs within the scene. Furthermore there is a real need to correct the dark (e.g. shadows) as well as the bright (e.g. overexposed) pixels in the scene in order to enhance colour base target detection.

Some compression functions, such as that shown in Figure 5-2(a) and (b)), only enhance pixels in the dark region. It is proposed here to use a two way bidirectional Pixel-Wise Non-Linear Transfer Function (PWNLTF) for enhancing pixels in the dark and overexposed areas. The form of the proposed transfer function is shown in Figure 5-2(d). Figure 5-2(c) illustrates the NLTF curve of NWIS method with different weighted parameter α and β (note here in Figure 5-2(c) some of the NLTF curves are overlapping each other).



(a)



(b)

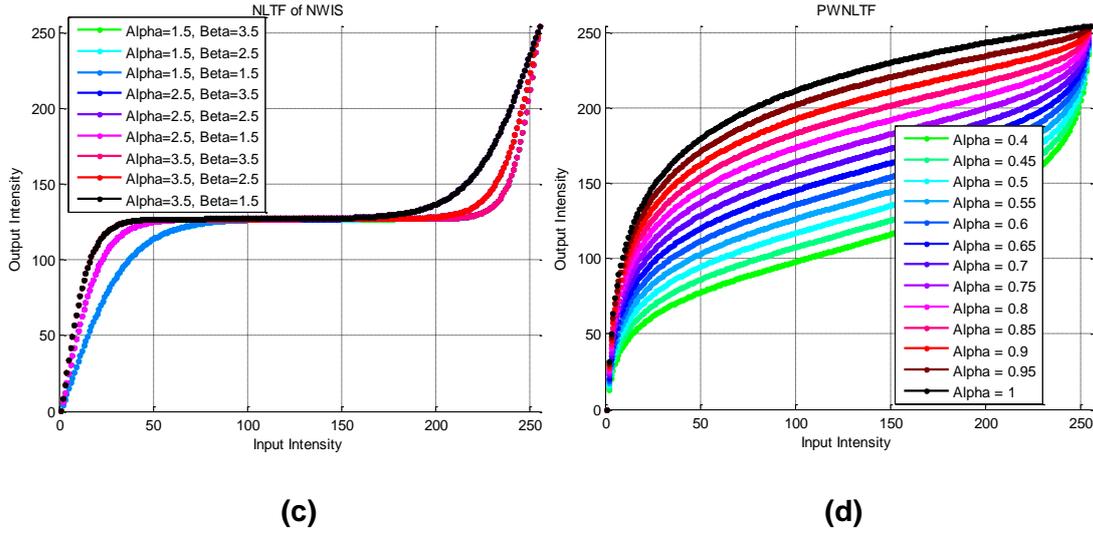


Figure 5-2: Illustrates different non-linear transfer function plots. (a) AINDANE method with weighted parameter “z” (b) IRN transfer function curve with parameter vmin (c) NWIS curve with α and β parameters (d) PWNLTF curve with weighted parameter α .

5.2 Bidirectional Pixel Wise Non Linear Transfer Function

5.2.1 RGB to HSV

Unlike RGB colour space, HSV separates luminance or intensity image (V) from chrominance or colour information of the image (H and S). The RGB image can be converted into HSV image:

$$H = \left\{ \frac{\frac{1}{2}[(R-G) + (R-G)]}{(R-G)^2 + (R-B)(G-B)} \right\}$$

where, (5.1)

$$H = \left\{ \begin{array}{ll} H, & \text{if } B \leq G, \\ 360 - H, & \text{if } B > G \end{array} \right\}$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \quad (5.2)$$

$$V = \max(R, G, B) \quad (5.3)$$

where H, S, V are Hue, Saturation and Value component of HSV colour space respectively.

5.2.2 Pixel Wise Non Linear Transfer Function (PWNLTF)

Mid tone dynamic compression T can be written in the form of:

$$I_{(x,y)} = T_{(x,y)}(x_{(x,y)}) \quad (5.4)$$

$$I_{(x,y)} = (x_{(x,y)})^T \quad (5.5)$$

where $x_{(x,y)}$ is input intensity value and $I_{(x,y)}$ is the output intensity value after it is transformed by transformation function $T_{(x,y)}$ at pixel location (x, y) , T is the power of transformation (Arigela & Asari, 2013). If p is $0 < T < 1$ then it amplifies the input value; if $T > 1$ then it compresses the data.

The form of the transformation has been suggested to be logarithmic as it is the way that human's visual perception works according to the Weber-Fechner law (Land & McCann, 1971; Hurlbert, 1989). Thus the function to be utilised for dynamic compression in this work is proposed to be based on two logarithmic functions:

General form:
$$I_{enh(x,y)} = \log(I_{(x,y)}) - \log(Z - I_{(x,y)}) \quad (5.6)$$

Specific form:
$$I_{enh(x,y)} = \alpha_{(x,y)}(\log(I_{(x,y)})) - (1 - \alpha_{(x,y)})(\log(Z - I_{(x,y)})) \quad (5.7)$$

where $I_{(x,y)}$ is the intensity at (x, y) , $Z=256$ for 8-bit images, α is a scale factor which depends on the illumination at the image pixel (x, y) . Examples of these transfer function is shown in Figure 5-3. The α parameter can be extracted from the scene:

$$\alpha_{(x,y)} = \left(\frac{I_{(x,y)}}{255} \right)^{(-c)} + \beta \quad (5.8)$$

where β is an enhancement dependent parameter for the whole frame of image data and c is a constant which is found from experiment (see Appendix B). Note that $I_{enh,(x,y)}$ is evaluated for every pixel through the pixel-wise parameter α .

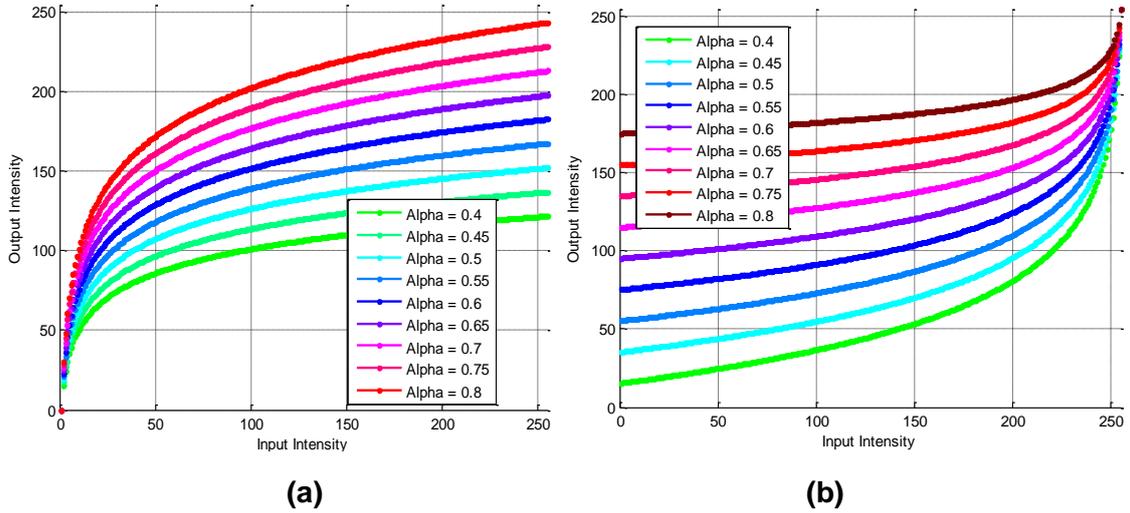


Figure 5-3: Outlines the bidirectional NLTF proposed in this work (a) first half of PWNLTF which is controlling dark part of images (b) Second half of PWNLTF which is controlling bright part of images.

β is the enhancement factor for a video frame and it is designed to maintain the mean intensity of the video clips to a constant over the video sequence. β is set at 0.6 for the first frame ($t+1$) of data, and subsequent β value can be computed by iteration process until the mean of the $I_{enh,(x,y)}$ in the next frame is close to the that of previous frame, or close to a user selected desirable $I_{enh,(x,y)}$. The flow chart for computing β is depicted in Algorithm 1.

Algorithm 1: Calculation of β

Frame $t=0$

Calculate mean of $I_{enh}(t=0)$, Initialise β (random value)

Next frame $t+1$

use $\beta=0.6$

while

 compute PWNLTF with β

 compare mean ($I_{enh}(t+1)$) with mean ($I_{enh}(t)$) or desirable I_{enh} mean by Euclidean Distance

if distance is large

 increase β by larger factor

elseif distance is small

 increase β by smaller factor

end

repeat

until (mean($I_{enh}(t+1)$) = mean($I_{enh}(t)$) or desirable I_{enh} mean)

This is possible that the β parameter can also be utilised to synchronise image quality across the CCTV network. It is postulated that the β parameter of other camera view can be adjusted such that the $I_{enh,(x,y)}$ in this view is made to synchronise with that in the reference camera view. Further experiment in this aspect will be studied in the future work.

The α parameter has direct effects on the dynamic range compression and also has strong influence on the image brightness (intensity). Figure 5-4 shows examples of the effects for different values of α and β in the PWNLTF. Figure 5-4(a) is raw image, Figure 5-4(b) left: processed image with PWNLTF and right: the PWNLTF curve with $\beta=0.8$ for all pixels. It is seen that the PWNLTF

compress high intensity pixels (e.g. flooring near to the exit) to low, and at the same time it brightens up the dark region (e.g. flooring near to the bottom of image). When the β is increased to 0.9 and 1 as shown in Figure 5-4(c) and Figure 5-4(d) respectively, the dynamic range compression is increased and the NLTF is shifted up accordingly.

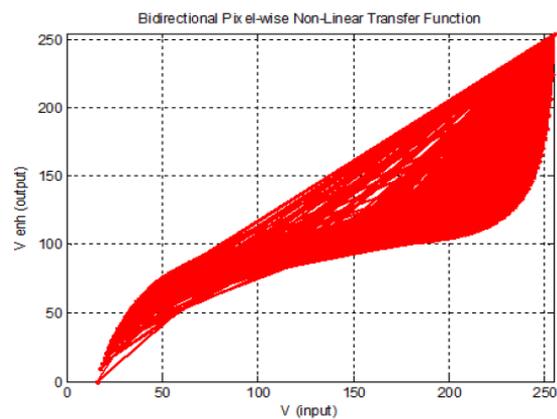
Like many other techniques, this methodology is quite dependent on the intensity profile of the very first initial frame data. Again, this issue will be further investigated in the future work.

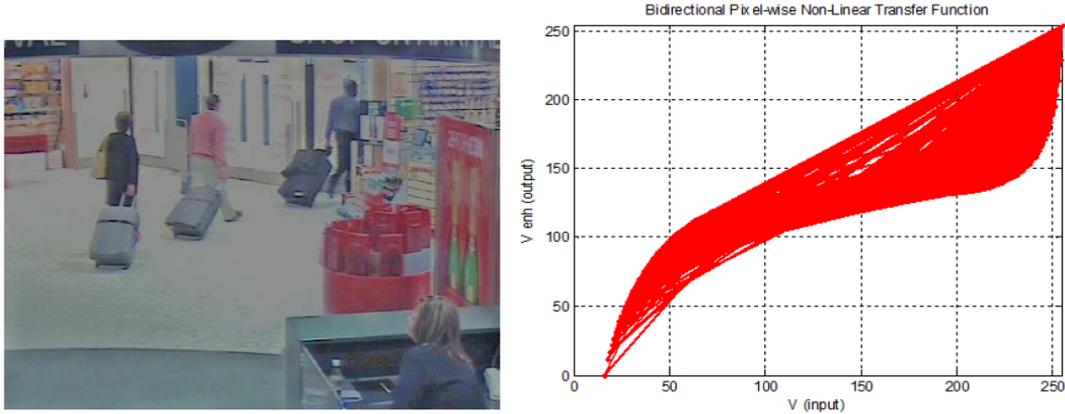


(a)

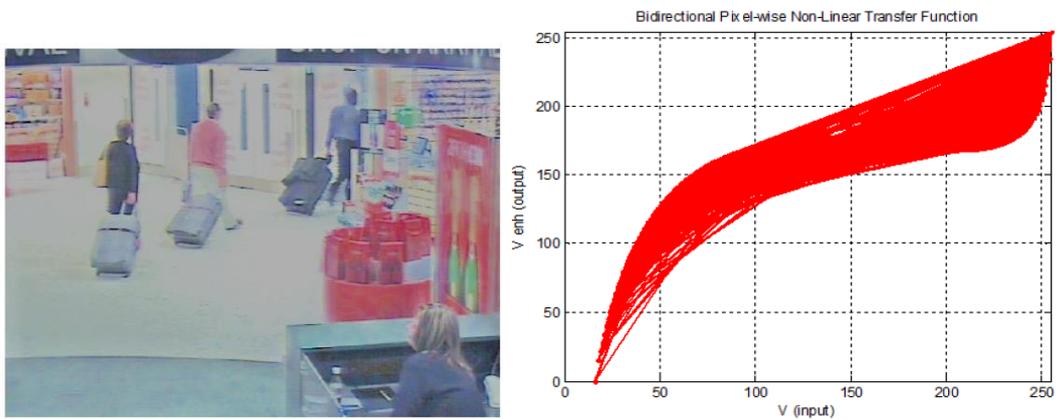


(b)





(c)



(d)

Figure 5-4: Shows the effect of α and β (a) raw image (b) left: processed image with PWNLTF method right: its corresponding curve, where $\beta = 0.8$, similarly in (c) and (d) but this time $\beta = 0.9$ and $\beta = 1$ respectively.

The V_{enh} intensity map can be normalised to:

$$I_{enh}' = \frac{I_{enh} - I_{enh, \min}}{I_{enh, \max} - I_{enh, \min}} \quad (5.9)$$

Where $I_{enh, \max} = \max(I_{enh})$ and $I_{enh, \min} = \min(I_{enh})$.

To avoid having pixels with high saturation the high contrast point can be found using:

$$HC_{(x,y)} = \frac{I_{raw} - I_{raw(avg)}}{I_{raw(avg)}}, \quad HC_{(x,y)} = I_{enh}'_{avg} \quad (5.10)$$

where HC is the high contrast point in the image, I_{avg} is the mean of V for the image frame. All pixels in the image with I larger than 5% of HC can be replaced by the mean (I_{enh}').

5.3 Centre Surround Luminance [CSL]

In the next stage of the work flow is to enhance the intensity V further using a centre-surround type of CC to help compensate colour degradation due to illumination artefacts. This part of work follows the algorithm proposed by Tao (Tao, *et al.* 2005).

The luminance can be obtained through a 2D discrete Gaussian (G) function as shown in equation (5.11):

$$I_{conv}(x, y) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} I(m, n) G(m+x, n+y) \quad (5.11)$$

G is the 2D Gaussian at pixel location (x,y), M and N are the dimension of the image:

$$G(x,y) = q.e^{-\left(\frac{x^2+y^2}{c^2}\right)} \quad (5.12)$$

where c is the surround neighbourhood constant (values of the neighbourhood constant c (10~20)). Note here that the equation (5.11) uses different surround constant c as compare to equation (4.4).

The q is the normalization constant computed via $\iint q.e^{-\left(\frac{x^2+y^2}{c^2}\right)} dx dy = 1$.

$$I_{enh}''(x, y) = I_{enh}'(x, y)^{E(x,y)} \quad (5.13)$$

where the exponent E(x,y) is defined by:

$$E(x, y) = \left(\frac{I_{conv}(x, y)}{I(x, y)} \right)^P \quad (5.14)$$

where $I_{(x,y)}$ is the intensity of the input at (x,y) , i.e. the I is V component of HSV and the $I_{conv(x,y)}$ is the centre-surround luminance evaluated using larger values of the neighbourhood constant as shown in equation (5.11). The $I_{enh''(x,y)}$ is the luminance after contrast enhancement. The exponent P is chosen to be a function of the global standard deviation in the $I_{(x,y)}$ which measures the extremeness of the intensity map. The exact value of P is determined by experiments and it can be scene dependent. Typical values of P is (0.5~0.75) for the dataset utilised in this study. Once the enhanced V component that is $I_{enh''}$ is computed by using equation (5.13), it is combined with H and S components in HSV space. Hence the CC transformed RGB colour space image is obtained by converting HSV image back to RGB space.

5.4 Effectiveness of PWNLTF+CSL vs other NLTF algorithms

It is known that previous work such as AINDANE (Tao & Asari, 2005) IRN (Tao, *et al.* 2005) and NWIS (Asari, *et al.* 2006) algorithms are not capable to handle local variable illumination scenarios. In Figure 5-5 it is shown how effective is the work developed here, i.e., the PWNLTF+CSL, to compare with that of the reported NLTF CC algorithms.

Figure 5-5 shows 4 frames of video clip extracted from camera 1 view of the iLIDS data set to highlight the effects of images under variable illuminations. Figure 5-5(a)-(e) are the images of the Raw data, after processed by AINDANE, IRN, NWIS and PWNLTF respectively. It can be observe that the images after processed by other NLTF methods cannot maintain colour constancy. In contrast, the image data that processed by the PWNLTF+CSL algorithm has shown rather good colour constancy throughout the whole video sequence.

This result is further evidenced from the RGB plot as shown in Figure 5-6. The RGB attributes of the Target (red box) in camera 1 after processed by all algorithms are shown in Figure 5-6(a)-(c). Note that the images that have been

processed by other NLTF algorithms have not been able to compensate the colour of the target due to variable illuminations. Likewise Figure 5-6 (d) to (f) depicts the Standard Deviation (SD) of the R, G and B bands respectively for the 10 frames ROI (red box) data.

The very small SD in the developed PWNLTF+CSL algorithm further demonstrate the better colour constancy capability in proposed methodology.

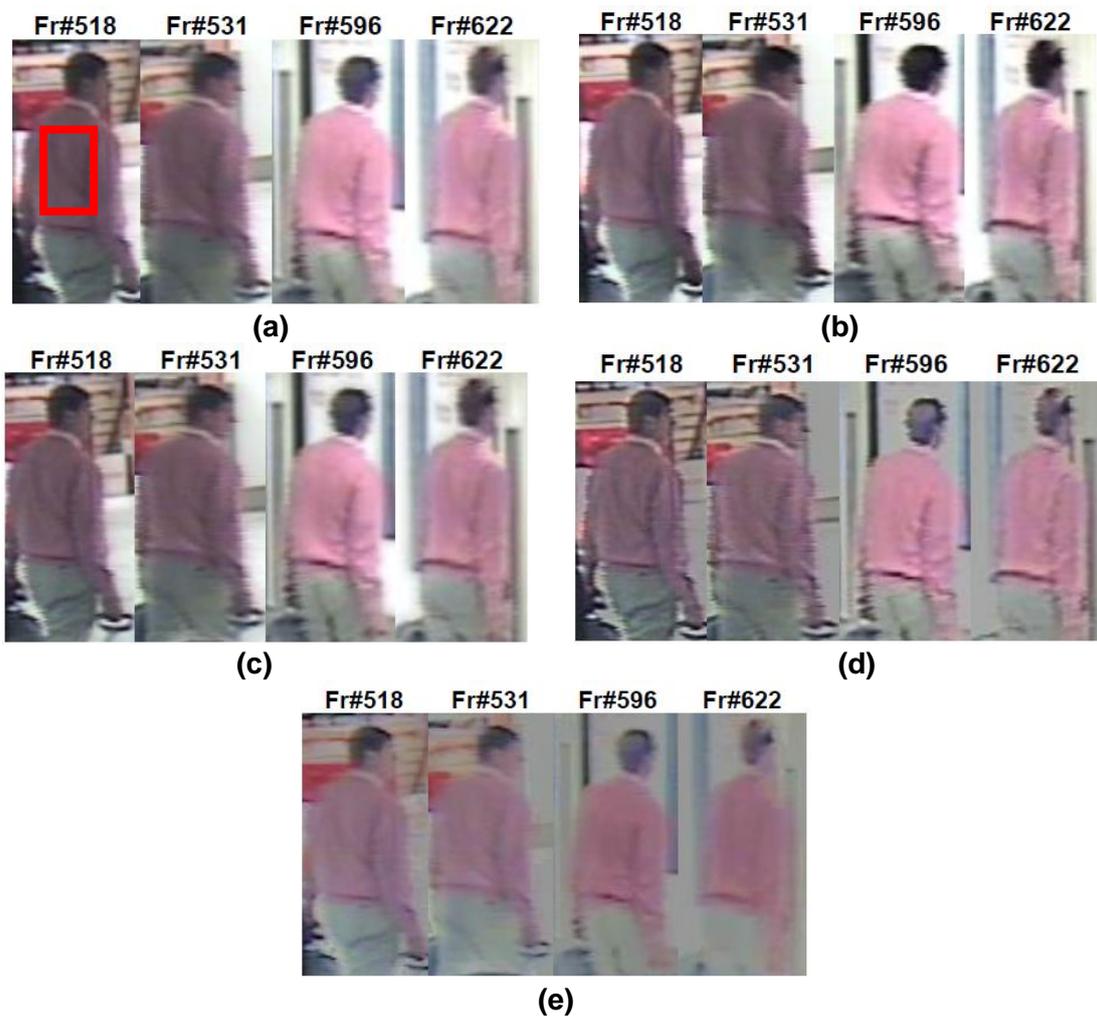


Figure 5-5: Shows the comparison of PWNLTF with other NLTF algorithms on Target (red box) in camera 1, frame numbers 518 and 531 are the entry point of target where illumination is weak and frame numbers 596 and 622 are middle and exit point of target where illumination is strong. (a) to (e) are Raw, AINDANE, IRN, NWIS and PWNLTF+CSL respectively.

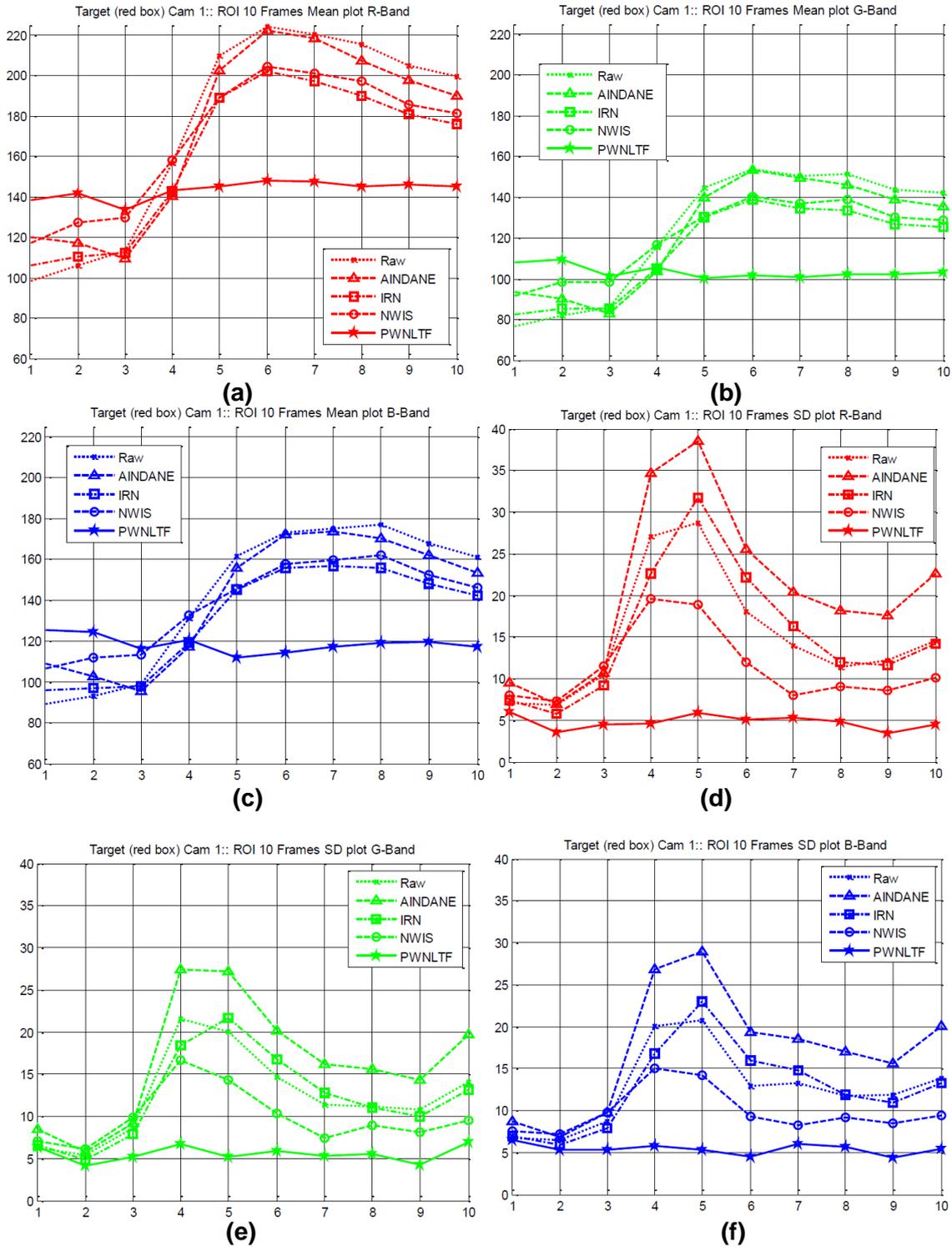


Figure 5-6: Highlights the RGB attributes of the ROI for 10 frames of Target (red box) in camera 1 after processed by PWNLTF+CSL and other NLTF algorithms as shown in Figure 5-5, (a) to (c) are R, G, and B bands respectively. Note that the images that have been processed by other NLTF algorithms are found not capable to rectify colour non-uniformity due to variable illumination conditions,

(d) to (f) depicts the Standard Deviation (SD) of the R, G and B attributes respectively of the ROI for 10 frames, this SD gives the indication that inside the ROI colour losses its constancy as target moves in different illumination regions.

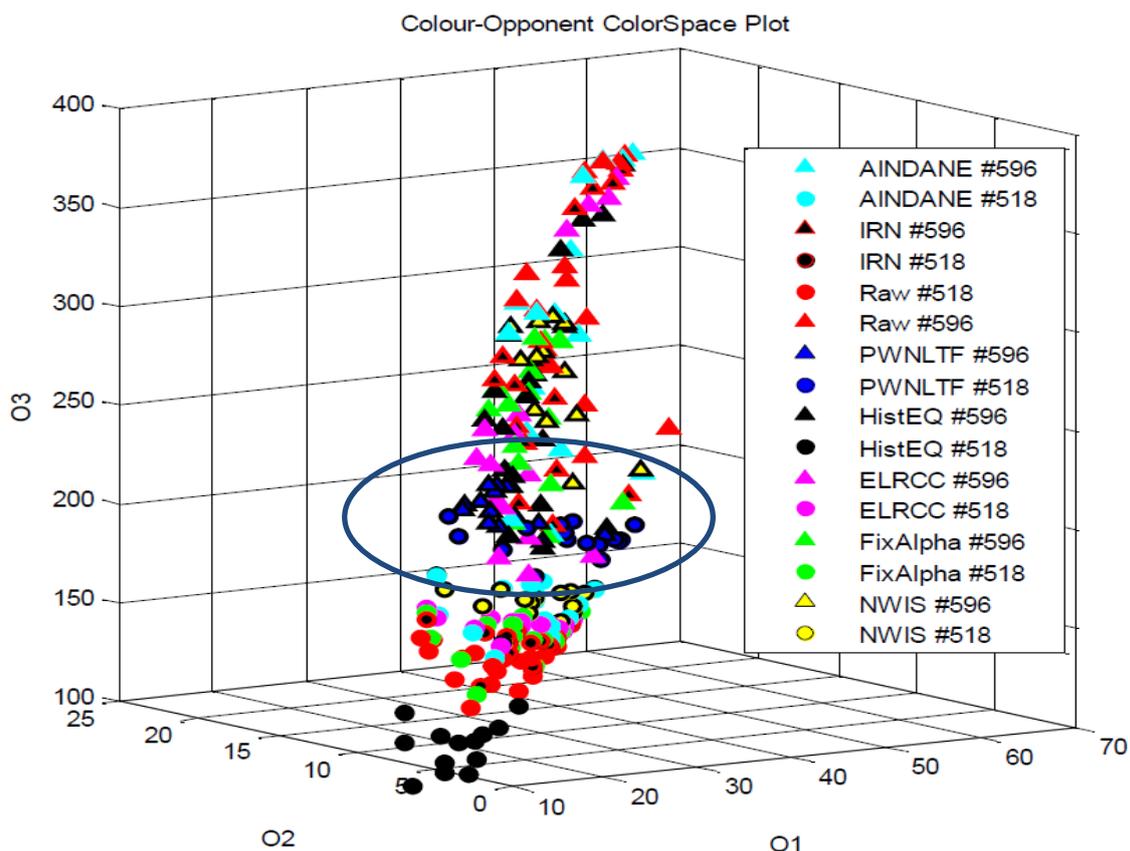


Figure 5-7: Shows the effectiveness of CC for various algorithms in the colour opponent colour space. The circle and triangle plots represent data extracted from the first (ie frame #518) and last (frame#596) frame of images as shown in Figure 5-5.

The effectiveness of colour constancy in the proposed algorithm can also be seen in the colour opponent space. Figure 5-7 shows the scatter plot of the opponency of the target after processed by various CC algorithms. The circle and triangle scatters represent 16 data points of target extracted from the first (frame #518) and last (frame# 596) frame of data as presented in Figure 5-5. It can be observed that only the plots in blue, i.e. the data after processed by the proposed PWNLTF+CSL method, exhibits almost the same opponency

throughout the video sequence. This shows the excellent colour constancy of the proposed CC algorithm with performance far better than all other reported CC methods studied in this work.

5.5 Grey Edge Algorithm

This is the third stage of the EDTCC work flow.

Grey Edge is a method which is capable to convert the spectral irradiance of an unknown light source into another light source with known spectral characteristics. The principle of the algorithm is the assumption that the average reflectance difference of all pixels in the scene is achromatic (Weijer, *et al.* 2007). The average of the colour derivatives, usually computed via the Minkowski norm, represents the characteristics of the illumination source:

$$e_c = \left(\int \left| \frac{\partial^n I_c(\mathbf{x})}{\partial \mathbf{x}^n} \right|^P d\mathbf{x} \right)^{1/P} \quad (5.15)$$

where e_c is estimated spectral characteristics of the illumination source for channel c , $c = [R, G, B]$ for broadband imaging, P denotes the Minkowski norm and n is Grey-Edge order.

The e_c is estimated for two sets of data and they are then normalised with a diagonal matrix which will then be used for rescaling the luminance of the other data set:

$$\begin{aligned} I_{r_{ColourCorrection}}(x) &= \frac{1}{\sqrt{3}} \frac{Ir(x)}{e_r} \\ I_{g_{ColourCorrection}}(x) &= \frac{1}{\sqrt{3}} \frac{Ig(x)}{e_g} \\ I_{b_{ColourCorrection}}(x) &= \frac{1}{\sqrt{3}} \frac{Ib(x)}{e_b} \end{aligned} \quad (5.16)$$

Where, $I_{r_{ColourCorrection}}$, $I_{g_{ColourCorrection}}$, $I_{b_{ColourCorrection}}$ are the colour corrected R,G,B bands of an image, e_r, e_g, e_b is the estimated the colour for each R,G,B band

respectively, factor $\sqrt{3}$ is the normalisation constant and (X) is the (x,y) coordinates of an image.

For convenience the processing algorithm up to this point, i.e., PWNLTF+CSL+grey edge is termed as ETDCC.

5.6 Colour invariant descriptors

This is the last step of the ETDCC algorithm work flow.

A range of colour invariant descriptors such as the Sum Feature(Soori, *et al.* 2011), L1 L2 L3 descriptors (Gevers & Smeulders, 1999), C1 C2 C3 descriptors (Gevers & Smeulders, 1999), the reduced rgb colour feature (Gevers & Smeulders, 1999) and O1 O2 O3 colour opponent feature (Gijsenij & Gevers, 2011) have been individually tested together with the enhanced luminance output obtained from the PWNLTF+CSL+Grey Edge (ETDCC) procedure, to formulate the ultimate ETDCC algorithm.

The final form of the ETDCC is targeted to achieve the following two important objectives:

- A. The ability to maintain colour constancy irrespective of illumination conditions and camera calibration issues.
- B. The ability to enhance target detections from multi-camera CCTV network systems.

In the following sections the formulation of the ultimate ETDCC and the colour constancy ability as well as the target detection performance of this algorithm, are presented.

5.7 Ultimate ETDCC formulation and test results

5.7.1 Targets Representations

Ten subjects with various coloured clothing have been chosen as targets in this chapter, and typical RGB images of these targets are shown in Figure 5-8.



Figure 5-8 : Shows the representative pictures of targets exploited in this chapter: left to right T1 to T10 respectively.

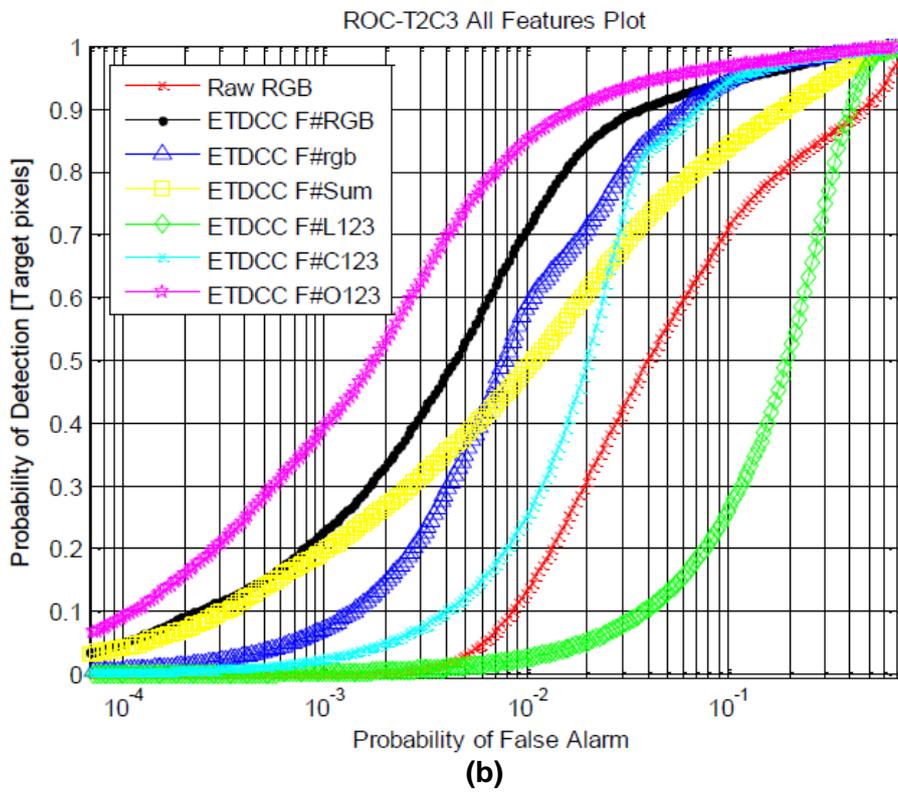
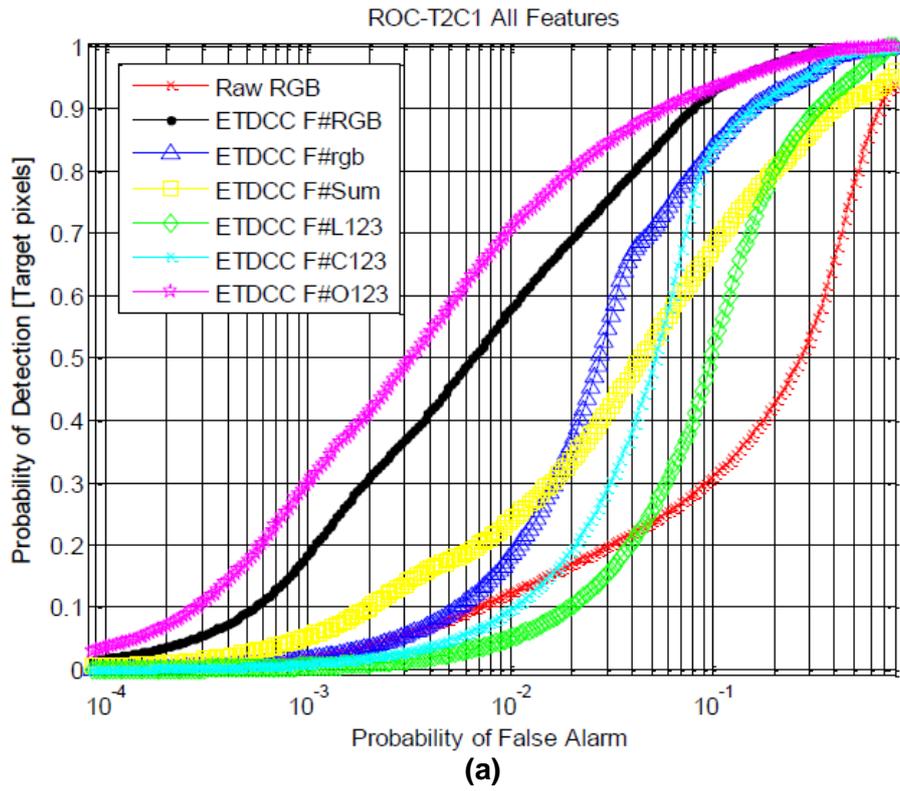
5.7.2 Colour Descriptors

Six different colour descriptors have been utilised:

1. RGB Feature,
2. Sum Feature (Soori, *et al.* 2011),
3. L1 L2 L3 descriptors (Gevers & Smeulders, 1999),
4. C1 C2 C3 descriptors (Gevers & Smeulders, 1999),
5. Reduced rgb colour feature (Gevers & Smeulders, 1999); and
6. O1O2 O3 colour opponent feature (Gijzenij & Gevers, 2011).

Figure 5-9 shows the averaged results of 10 frames of images for the detection of target T2 in all 3 camera views using these six different colour descriptors. Note that some descriptors, such as the L1L2L3, have exhibited very poor colour invariance property thus giving rather poor detection performances. The C1C2C3 and reduced rgb descriptor also gives quite variable results which can be seen from the ROC of the two camera views as shown in Figure 5-9 (a), (b) and (c).

Table 5-1 summarises the detection results in FOM metric for the experiments as presented in Figure 5-9. It is seen that the O1 O2 O3 colour opponent feature together with ETDC has exhibited the best detection performances over all other descriptors. Hence the O1 O2 O3 colour opponent descriptor feature has been selected to formulate the ultimate ETDC algorithms.



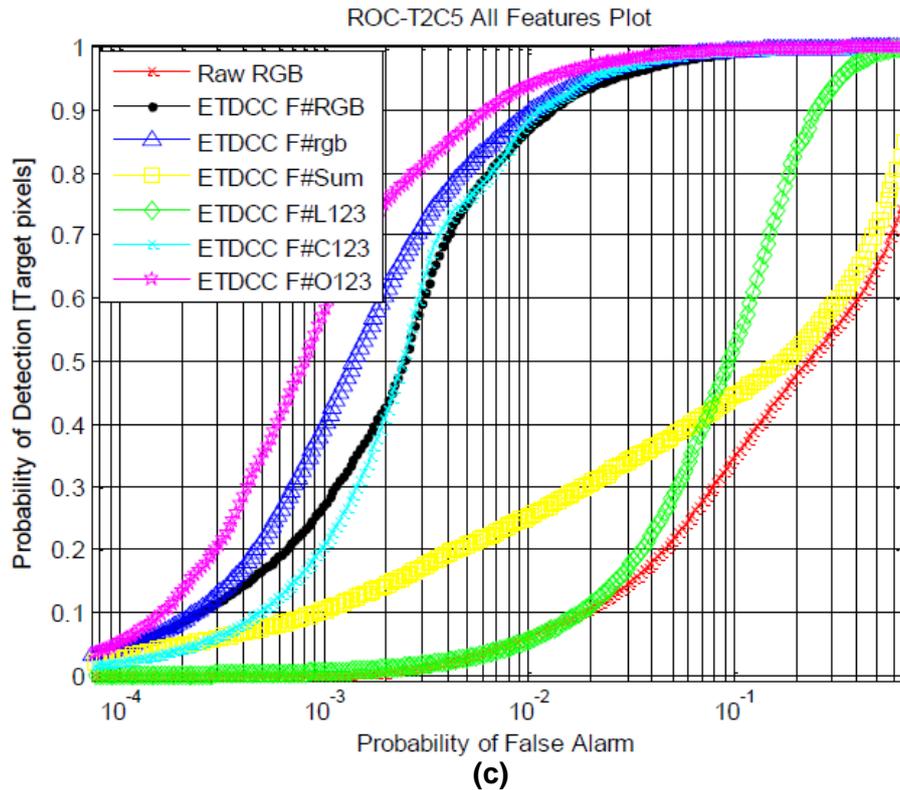


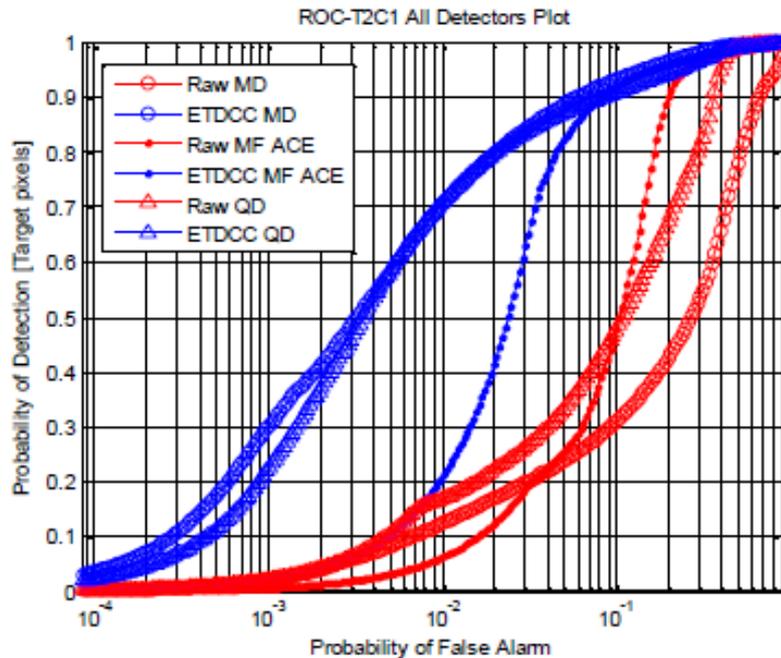
Figure 5-9: Shows the averaged detection ROC of T2 with respect various descriptor in (a) camera view #1 (b) camera view #3 and (c) camera view #5.

Table 5-1: Shows the efficiencies of different features utilised in this research study in FOM metric.

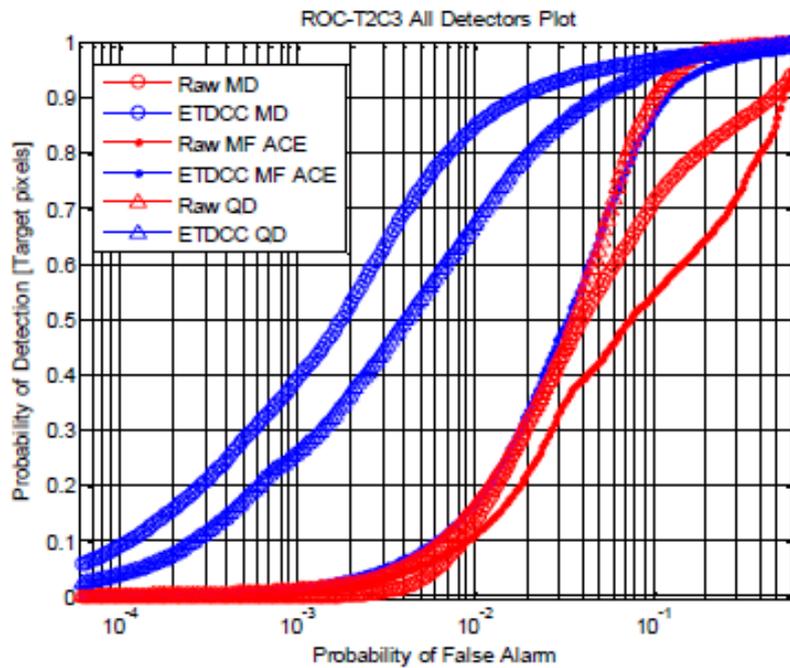
	Features	Camera 1	Camera 3	Camera 5	Mean FOM in %
1	O1 O2 O3	175.40	111.47	286.46	191.11
2	Sum	63.10	56.61	57.94	59.22
3	L1 L2 L3	20.40	-0.83	42.08	20.55
4	C1 C2 C3	52.54	32.96	233.86	106.45
5	rgb	75.14	56.61	260.20	130.65
6	RGB	146.64	83.73	241	157.12

5.7.3 Comparison of Detectors

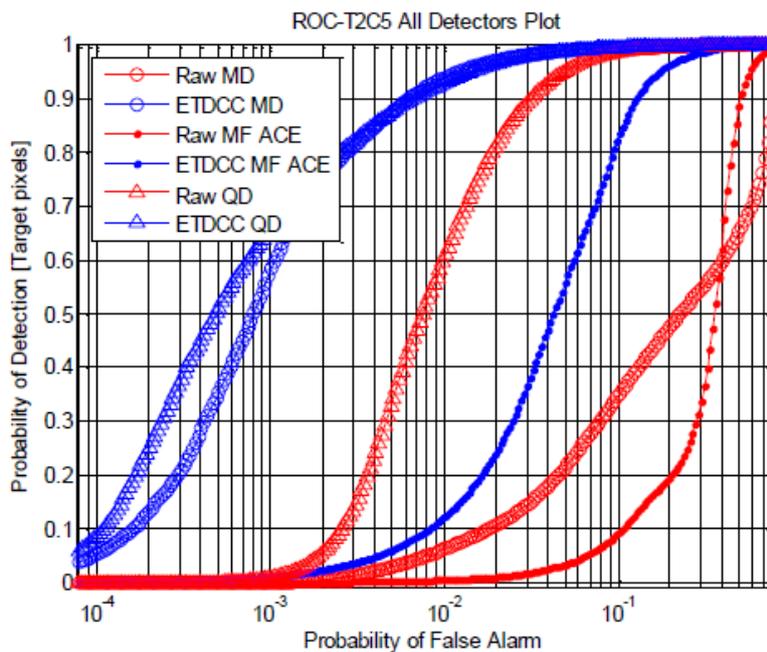
Figure 5-10 shows the ROC results for the detection of target T2 in all 3 camera views using 3 detectors namely: 1) Mahalanobis Discriminant (MD), 2) Quadratic Discriminant (QD) and 3) Adaptive Coherence Estimator (ACE), and compare them with that using the raw data, and also the processed data using the ultimate ETDC algorithm. The red traces denote the ROC obtained from the raw data in RGB feature, and the blue ones represent the detection results using the ETDC data. It is seen from Figure 5-10 and Table 5-2 that the detection performance amongst the three detectors varies significantly. This may be due to the limited colour channels in this RGB data set which may affect the conditions of the covariance in the colour opponent space. It is seen that the MD detectors together with the ETDC have exhibited much better detection performance than that of the QD and ACE.



(a)



(b)



(c)

Figure 5-10: Shows the detection ROC for target T2 (a) camera view #1 (b) camera view #3 and (c) camera view #5 using the 3 detectors MD (circle), QD (square) and MF ACE (triangle) for the raw data and that after processed by CC algorithm (the ultimate ETDCC).

Table 5-2: Tabulates the detection efficiencies of three different detectors that have been employed in this study.

	Features	Camera 1	Camera 3	Camera 5	Mean FOM In %
1	MD	175.4	111.47	286.46	191.11
2	MF ACE	48.32	34.10	162.02	81.48
3	QD	98.91	58.72	53.92	70.51

5.7.4 Detection Performances of Various CC Algorithms

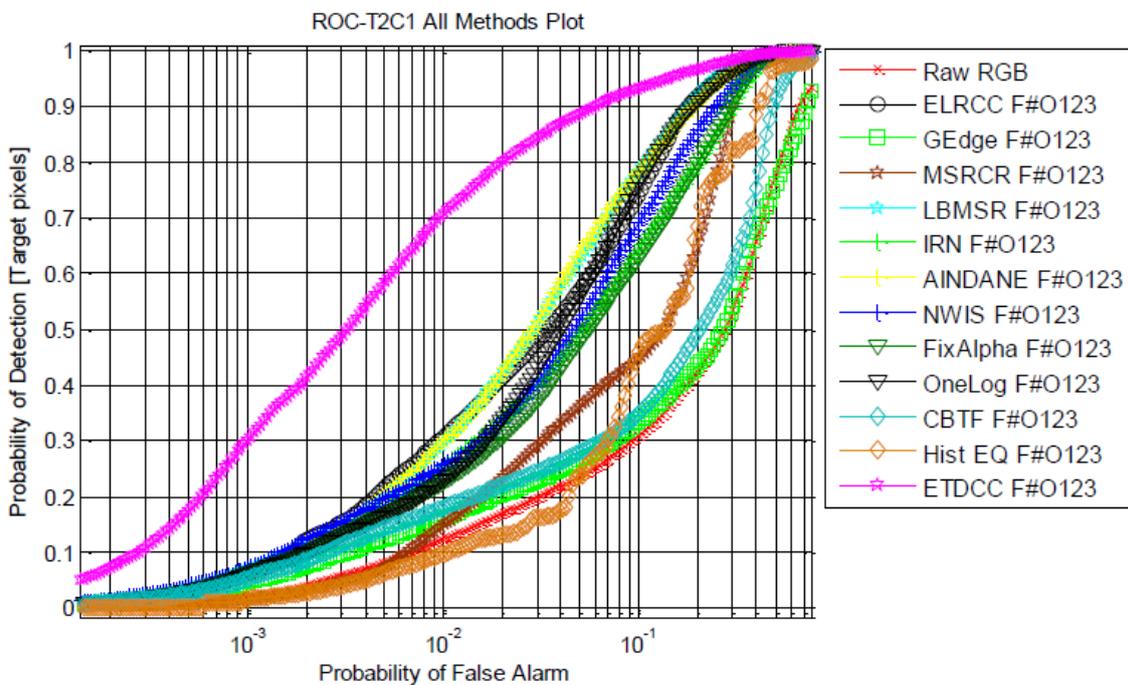
Twelve different CC algorithms reported in the literature have been utilised here to compare with the detection performance of the proposed ELRCC algorithm:

1. Enhanced Luminance Reflectance Colour Constancy (ELRCC) algorithm as defined in chapter 4 (Soori, *et al.* 2013),
2. Grey-Edge method (GEdge) (Weijer, *et al.* 2007),
3. Multi Scale Retinex with Colour Restoration method (denoted as MSRCR) (Rahman, *et al.* 1998),
4. Luminance Based Multi Scale Retinex algorithm (denoted as LBMSR) (Tao & Asari, 2003),
5. Illuminance-Reflectance Nonlinear algorithm (denoted as IRN) (Tao, *et al.* 2005),
6. Adaptive and Integrated Neighbourhood Dependent Approach for Nonlinear Enhancement method (denoted as AINDANE) (Tao & Asari, 2005),
7. Multilevel Windows Inverse Sigmoid function (denoted as NWIS) (Asari, *et al.* 2006),
8. The proposed PWLNTF dynamic compression but using fixed alpha (alpha=0.85) (denoted by FixAlpha),
9. Single logarithmic function as the dynamic compression (denoted as OneLog),
10. Cumulative Brightness Transfer Function method (denoted as CBTF) (Prosser, *et al.* 2008),
11. Histogram equalisation (denoted as Hist EQ) and

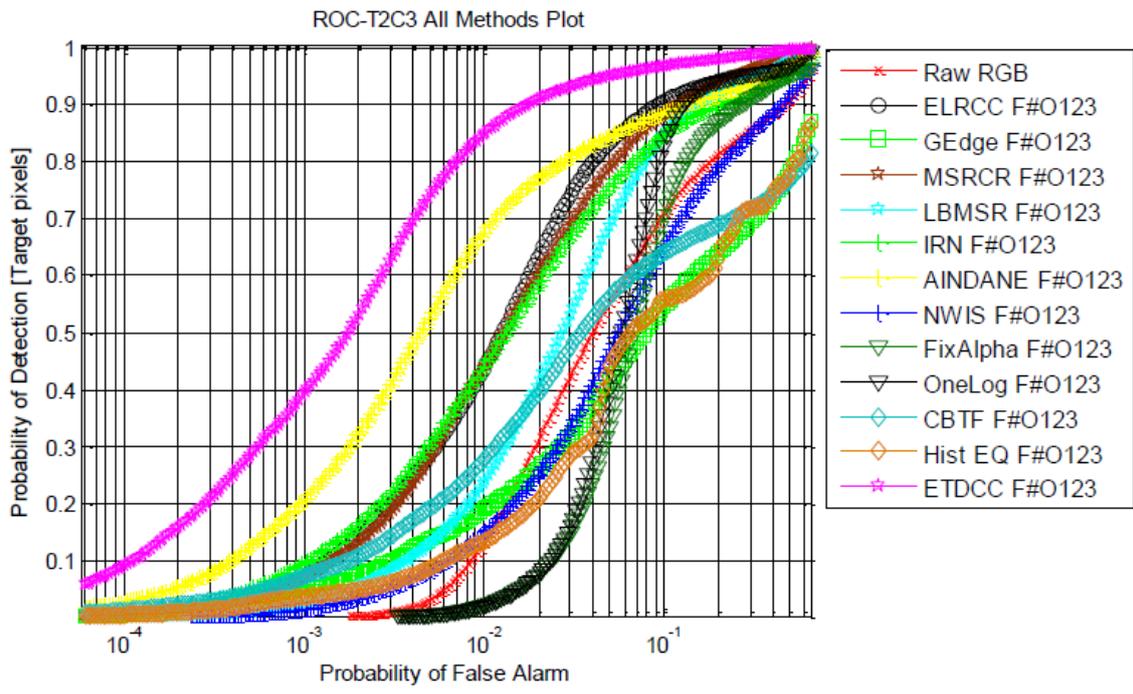
12. The proposed Pixel-wise Nonlinear Transfer function+CSL+Grey Edge: this is the ultimate ETDC algorithm (denoted by ETDC in the legend).

In all cases the O1O2O3 colour opponent feature has been employed throughout this experiment and the detection was performed using the MD classifier.

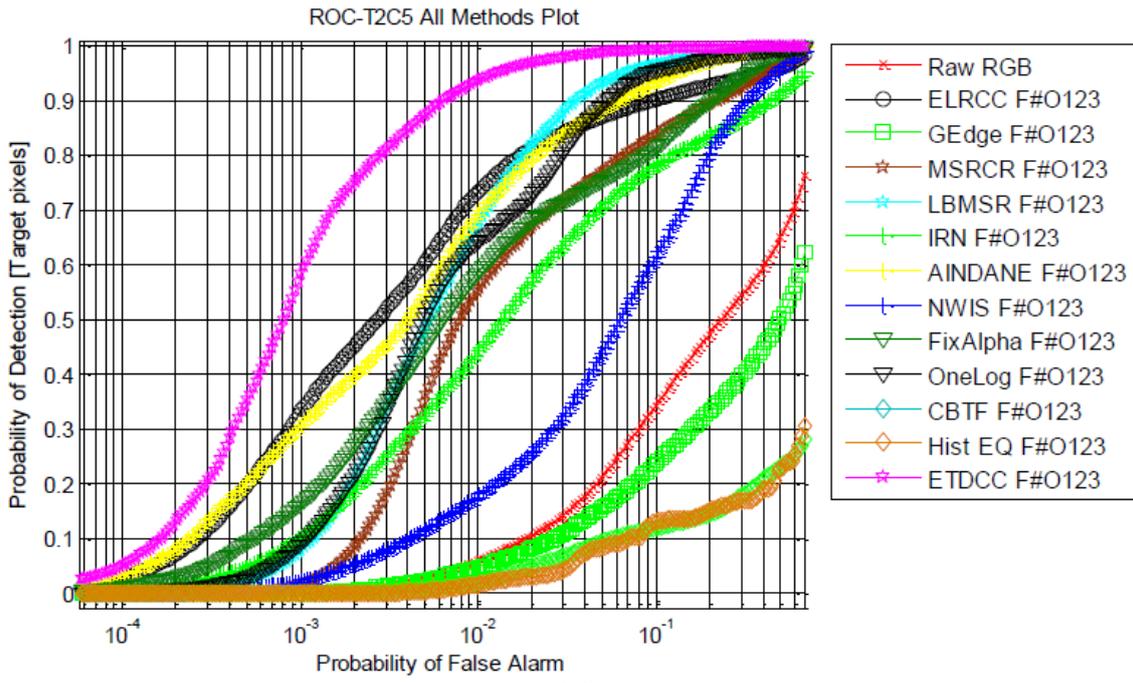
Figure 5-11 shows the detection results of target T2 for 3 different camera views after the image data is processed by the twelve different CC algorithms as mentioned above. It is quite clear that all other CC algorithms, apart from the proposed ETDC method, are found not capable to detect the target robustly. Note that some of the CC algorithms exhibit worse detection performances than that using the raw data. Much of the detection enhancement in the ETDC algorithm seems to be lying in the low probability of false alarm (PFA) region. Table 5-3 tabulates the FOM for the detection of Target T2 in 3 different views after the data is treated by these twelve different CC algorithms. It is quite clear that the proposed ETDC enhances target detections far better than all other CC methods reported in the literature.



(a)



(b)



(c)

Figure 5-11: shows the average of 10 detection results for target T2 in (a) camera view #1, (b) camera view #3, (c) camera view #5, after the video sequence is processed by twelve different CC algorithms. Note that all CC algorithms, except that processed by the ETDCC (magenta star) method, exhibit little or no consistence detection improvements with respected to that of the raw untreated data (in red x).

Table 5-3: Tabulates the FOM for the detection of Target T2 in 3 different views after the data is treated by twelve different CC algorithms.

	CC Methods	Camera 1	Camera 3	Camera 5	Mean FOM In %
1	ELRCC	82.95	45.73	213.45	114.04
2	GreyEdge	9.33	-9.23	-21.49	-7.13
3	MSRCR	31.22	42.94	138.59	70.92
4	LBMSR	86	20.65	182.16	96.27
5	IRN	66.92	41.95	130.69	79.85
6	AINDANE	86.47	72.33	212.15	123.65
7	NWIS	72.47	-2.43	68.42	46.15
8	FixAlpha	61.74	-12.97	164.89	71.22
9	OneLog	75.30	-6.38	177.18	82.03
10	CBTF	20.53	6.72	-62.15	-11.63
11	Hist Eq	16.57	-13.89	-63.41	-20.24
12	ETDCC	175.4	111.47	286.46	191.11

5.7.5 ETDCC & other CC algorithms: detection of 10 targets

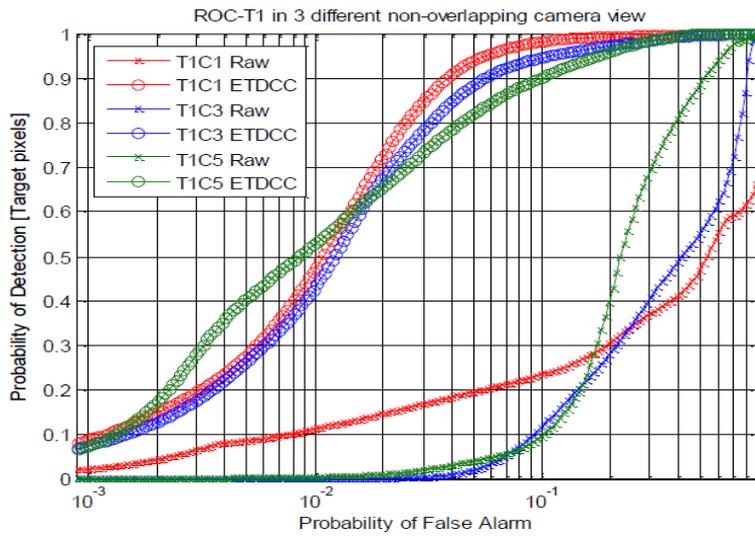
The capability of the proposed ETDCC algorithm is critically assessed here for the detection of 10 targets from the iLIDS data set. Figure 5-12 shows the typical RGB images of three camera views (view #1, view #3 and view #5) of target T1 with the raw data shown in the upper panel, the same data but after processed by ETDCC is shown in the lower panel. Note that the colour bleaching is very minimal after the data is processed by the ETDCC algorithm.



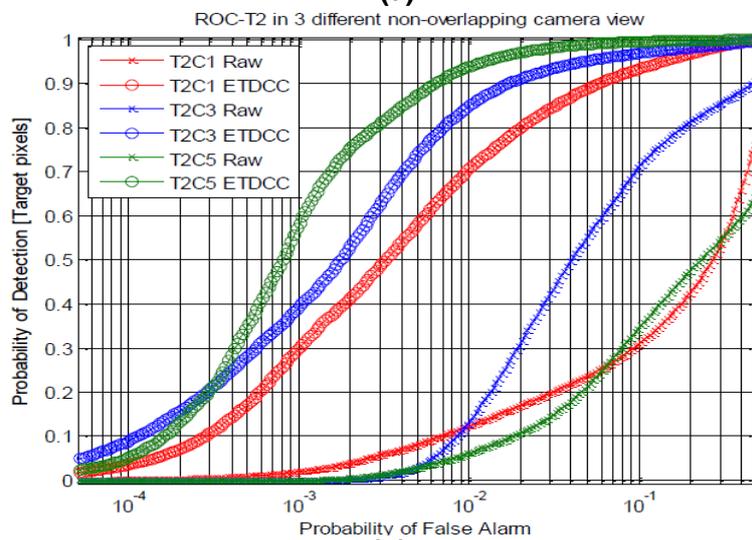
Figure 5-12: Shows representative images selected from the all 3 camera views of target T1 (left to right camera 1, 2 and 3 respectively). (a) Raw data. (b) After transformed by the ultimate ETDC algorithm.

Due to the large number of targets, the detection results of a subset of five (T1, T2, T4, T6 and T9) are presented here in Figure 5-13 (a) to (e). The ROCs are colour coded, depicting the results from camera views #1, #3 and #5 in red, blue and green respectively. It is quite clear that the ETDC processed data (in circle trace) exhibits consistently improved target detections over the raw data (in dot trace) for all camera views. The degree of enhancements can be indicated by the shift of the ROC curve to the right which can be measured more quantitatively from the area of the ROC within the PFA bounds of $[10^{-5}, 1]$.

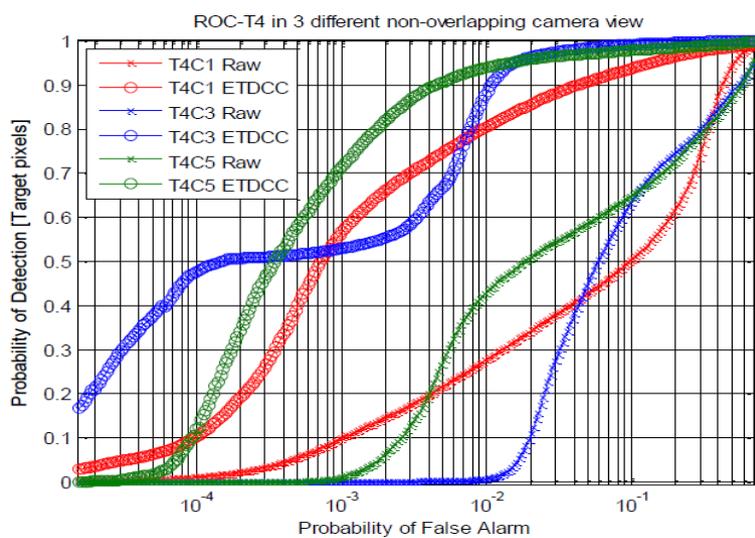
The enhancement of target detection after the data is processed by ETDC is computed for all 10 targets for all camera views using the FOM formulation according to equation (3.24). The result is tabulated in Table 5-4 which reveals an averaged of ~165% better detection for all these 10 targets has been resulted from the data after processed by the proposed ETDC algorithm.



(a)



(b)



(c)

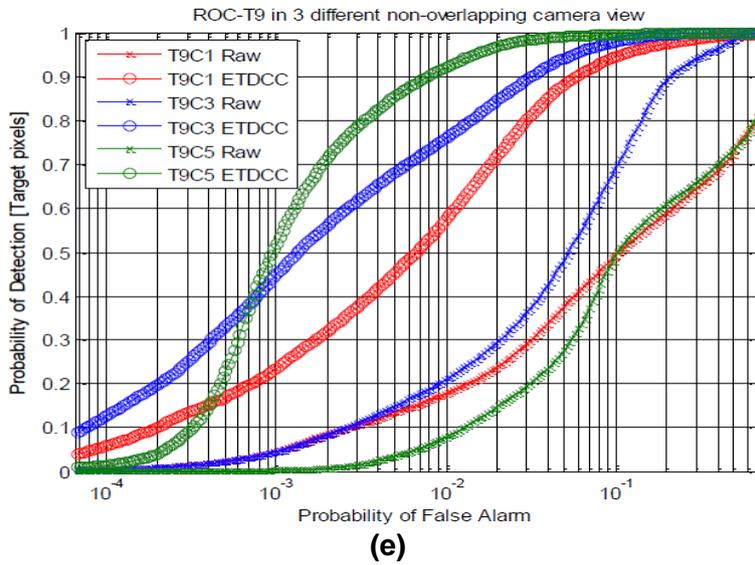
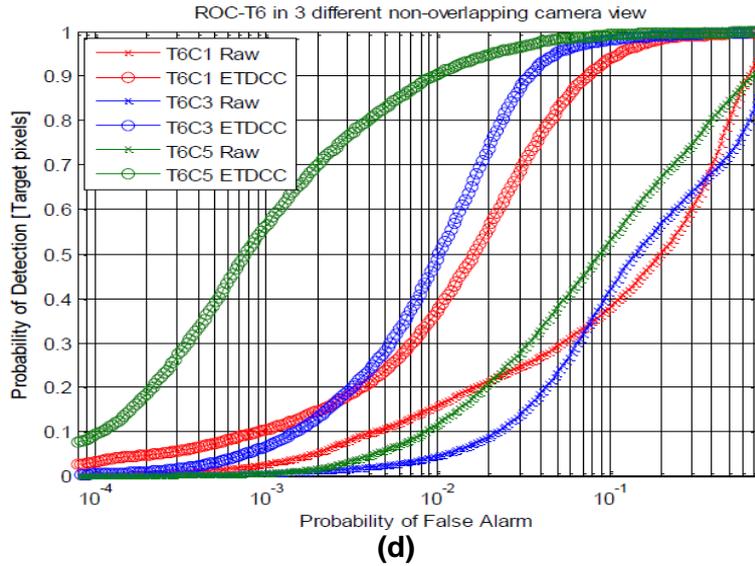


Figure 5-13: Shows the mean of the ROC results for the detection of targets (a) T1, (b) T2, (c) T4, (d) T6 and (e) T9. The ROC of the raw data and that after processed by ETDCC are presented in circle and cross respectively. The three different camera views of data are presented in colour code of red, blue and green for view #1, view #2 and view #3 respectively. Note that the circle plot ROC for all colour plots is substantially better than that of the cross in all cases. The colour opponent descriptor and the MD detector have been employed in all cases.

Table 5-4: Tabulates the FOM for the detection performance enhancement for the data processed by the proposed ETDCC for all 10 targets.

Targets	Camera 1	Camera 3	Camera 5	Mean FOM in%
T1	211.79	310.83	205.84	242.82
T2	175.4	111.47	286.46	191.11
T3	43.71	105.2	110.6	86.50
T4	109.11	205.82	113.29	142.74
T5	76.46	240.62	543.51	286.86
T6	92.09	139.74	175.89	135.91
T7	81.92	Target Absent	130.61	106.27
T8	79.18	249	266.78	198.32
T9	107.91	92.89	206.64	135.81
T10	79.55	Target Absent	144.74	112.15
Averaged	105.71	181.95	218.44	168.70

5.7.6 Discussion

This chapter details the formulation of the proposed ETDCC algorithm, and to verify its CC capability as well as target detection performance, within the context of other reported CC algorithms in the literature.

The main result of this chapter is presented in Figure 5-13 and Table 5-4, which highlights how the proposed ETDCC performs with respected to the detection of 10 targets from the iLIDS data set. The result gives the evidence that the CC processing can indeed reduce illumination artefact and to help restore the colour integrity of the scene to certain degrees. The averaged target detection enhancement for the ETDCC over the 10 target for the 3 camera views #1, #3 and #5 are found to be respectively ~105%, ~180% and ~215% better than that using the raw data. However, much enhancement has been seen from the detections in camera view #5, where the detection ranges is the longest amongst all other views. It is expected to be more difficult for target detection in

this view, because the background is strongly affected by the solar illumination through the window.

The average detection enhancement for all camera views over all 10 targets is found to be ~155% in FOM metric. In order to visualise the implication of this enhancement as the result of the CC treatment by the EDTCC algorithm, all the ROCs that presented in Figure 5-13, together with that of the rest of the other 5 targets, are summed and averaged as shown in Figure 5-14. It can be observed quite clearly from this Figure 5-14 that the ROC of the ETDCC processed data has shifted rather significantly to the lower PFA. At PD=0.5 the PFA reduction is found to be about two orders of magnitude after the data is treated by the proposed ETDCC.

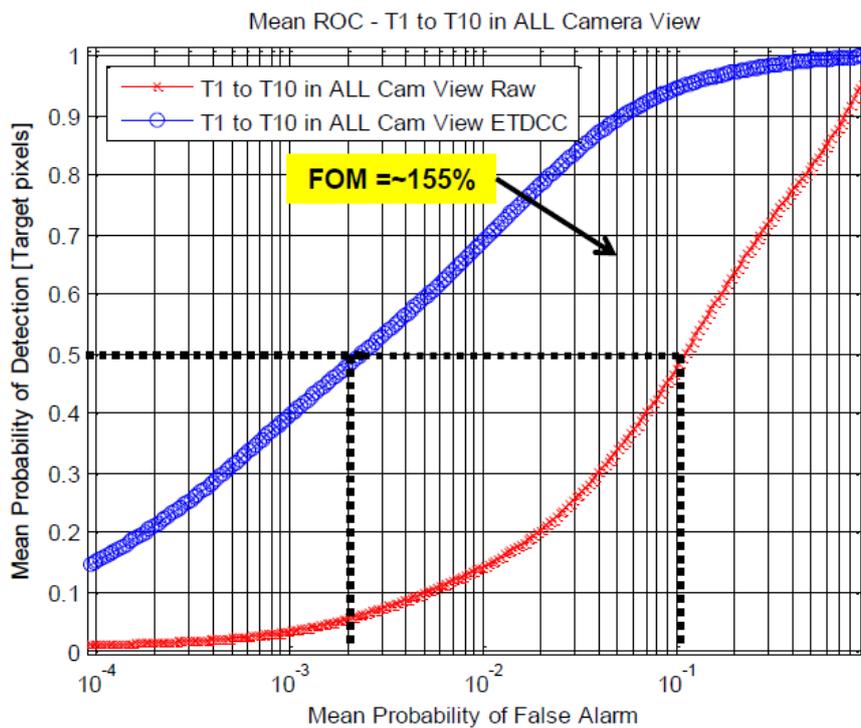


Figure 5-14: Shows the averaged ROC detection results for all 10 targets in all camera views, (red cross) ROC for raw data, (blue circle) ROC for after processing data with the ETDCC method. It can be seen that FOM of target enhancement by 155% FOM, implying for a PFA reduction by about two order of magnitude at PD=0.5.

Chapter 6: Discussions for all data sets [iLIDS, PETS, GTCC]

6.1 Targets

Various subjects with different coloured clothing have been chosen as targets in this chapter and typical RGB images of these targets are shown in Figure 6-1(a)&(b).



Figure 6-1: Shows the representative pictures of targets exploited in this chapter (a) left to right T1 to T12 respectively extracted from iLIDS dataset (b) left to right T1 to T4 respectively extracted from PETS 2007 dataset.

6.2 Limitation of colour invariant descriptors (rgb and C1C2C3)

Colour invariant feature such as reduced rgb and C1C2C3 as outlined in section 2.3.3.4 and section 2.3.3.3 respectively, have limitations to work for weak colours when the colour attributes are about the same in all 3 channels. Different shades of white, grey and black colours will be almost zero or infinity in these colour descriptors. This is illustrated in Figure 6-2 which outlines an image extracted from the GTCC dataset. The black/white and various shades of

grey colours becomes zero when the image is converted into reduced rgb (Figure 6-2(b)) and (Figure 6-2(c)) C1C2C3 colour space.

Another example can be seen in the iLIDS dataset as shown in Figure 6-3.

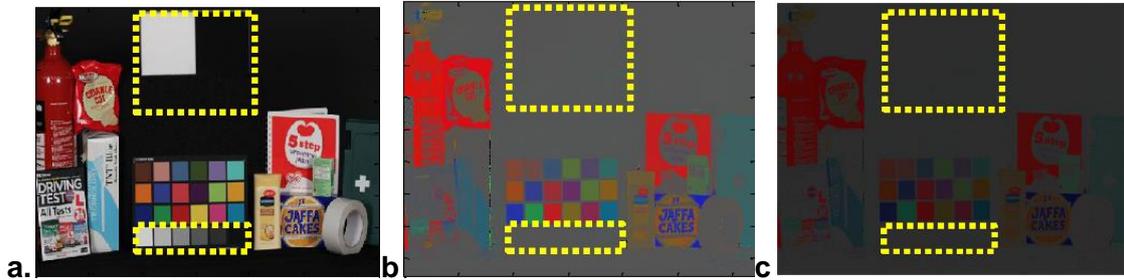


Figure 6-2: Highlights the drawback of some colour invariant descriptors (a) raw RGB image extracted from GTCC dataset (b) after converted image into reduced rgb colour space (c) after converted image into C1C2C3 colour space.

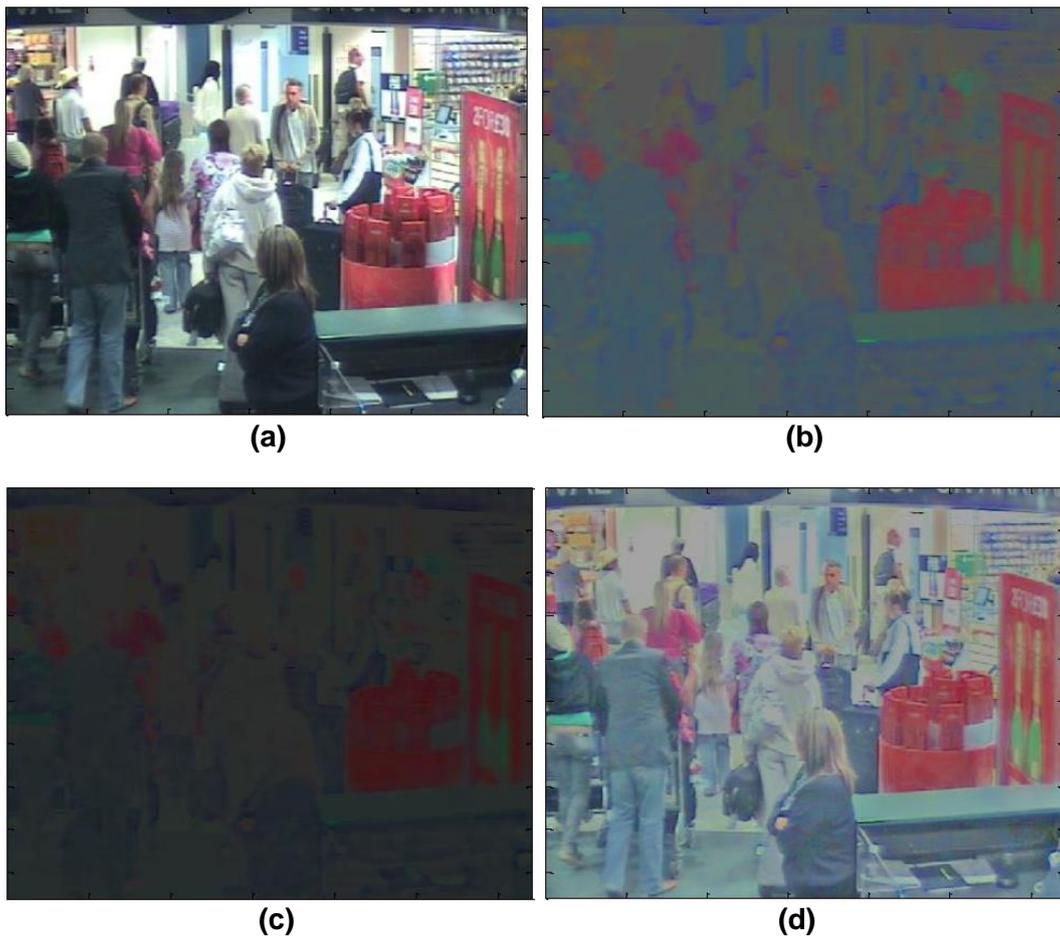


Figure 6-3: shows another examples of limitations in some colour invariant features (a) raw RGB image from iLIDS (b) after converted image into rgb colour

space (c) after converted into C1C2C3 colour space and (d) after processing the image using EDTCC.

The ineffectiveness of these colour invariant features has direct impacts on the target detections. Figure 6-4 shows the ROC of target 9 who has a light blue colour top. The detection performance for the data in the reduced rgb and C1C2C3 are quite poor with respect to EDTCC method which employs O1 O2 O3 colour feature. However, the detection performances for target 7 who was in sharp (strong) orange colour top shows almost the same result for all these 3 colour descriptors (see Figure 6-4).

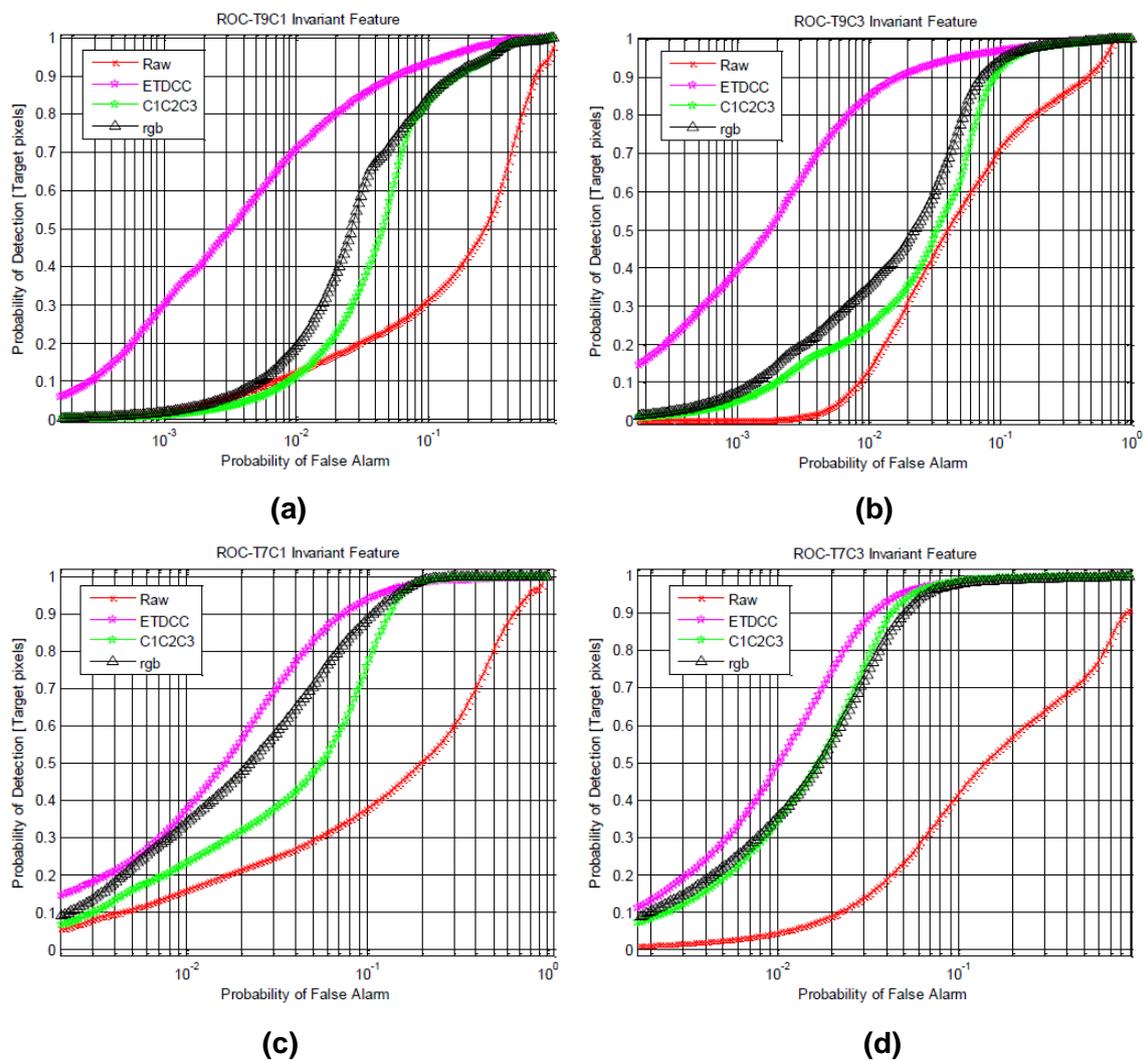


Figure 6-4: Demonstrate the impacts of limitations in colour invariant descriptors on the target detection performance: (a) and (b) ROC for weak coloured target 9

in camera 1 and camera 3 respectively, (c) and (d) ROC for strong coloured target 7 in camera 1 and camera 3 respectively.

6.3 Colour constancy over multiple camera views

In this section the colour constancy ability of various CC algorithm over all (3) camera views are presented as shown in Figure 6-5. The O1O2O3 chrominance of target 1 from camera views [1, 3, 5] after transformed by various CC algorithms are depicted in [circle, triangle, square] respectively. It can be seen that all tested methods, except for the one processed by the proposed ETDCC (in yellow plot), shows very good colour constancy over the 3 cameras views.

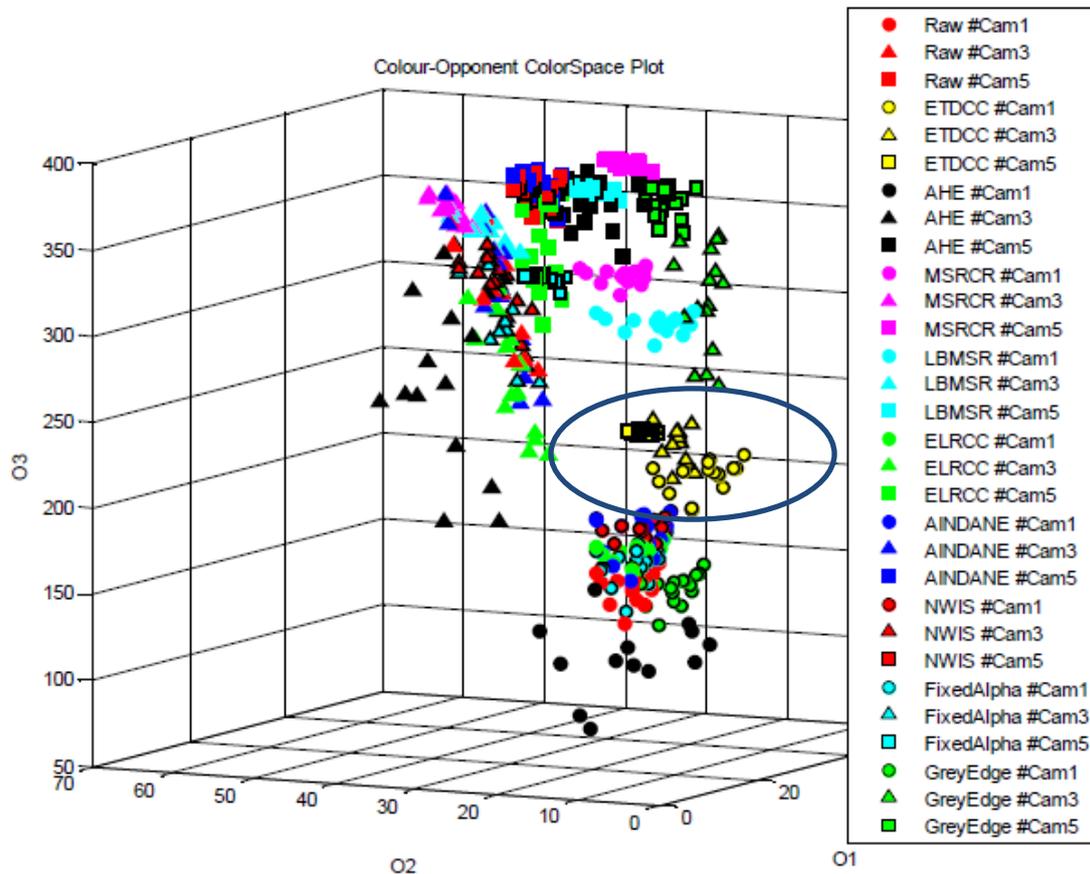


Figure 6-5: Shows the O1 O2 O3 chrominance of target 1 for all three camera views after processed by all CC algorithms utilised in this study. The plots for views [1, 3, 5] are in [circle, triangle, square] and note that only the proposed ETDCC (in yellow) exhibit good colour constancy (i.e. small spread).

6.4 iLIDS & PETS data sets: All CC algorithms

6.4.1 Colour Enhancements

Colour bleaching has been one of the main issues in almost all CC algorithms. Figure 6-6 highlights 3 views of a screen shot of target 6 extracted from the iLIDS data set after they are processed by various CC algorithms. Note that some CC algorithms could indeed enhance the colour and textural (contrast) but almost all of them fail to maintain the colour in a stable state (i.e. colour constancy). Figure 6-6(b)-(g) show the images after processed by the following CC algorithms: Histogram Equalisation (by colour and intensity), Average intensity, MSR, MSRCR and LBMSR methods respectively.

Figure 6-6(b) & (c) highlight the significant effect of colour bleach in target's t-shirt colour after the images are processed by the Histogram Equalisation method, showing the incapability of this algorithm to maintain colour constancy in variable illumination situations. The failure is mainly due to the histogram stretching in each colour channel which distorts the overall colour and in many cases it weakens the overall colour attributes (see Figure 6-6(b&c)). In Figure 6-6(d) it shows the result after the Average Int algorithm is employed as defined in equation (2.33). The images of the scene become darker because the algorithm compresses the image intensity significantly. The LBMSR result as shown in Figure 6-6(g) shows slightly better colour contrast as compare to other methods that presented in (b)-(f) because LBMSR evaluates luminance from each pixel rather than manipulating the colour attributes from the whole scene as in the conventional Retinex methods.

Figure 6-6(h)-(j) shows the images after processed by NLTF based methods and the result is seen to suffer from the same drawbacks as in the raw untreated data: poor colour constancy accompanying with strong colour bleach. This symptom is again mainly caused by using global parameters evaluated from the whole scene. In Figure 6-6(j), the NWIS compresses high intensity pixels through the NLTF (see Figure 5-2(c)) and exhibit better colour constancy as compare to IRN and AINDANE methods (see Table 6-4). The FixAlpha method which uses a fixed global parameter (Figure 6-6(k)) also is seen cannot

maintain CC efficiently. The OneLog method uses a non-linear transfer function to compress pixels only with low radiance and leaving the high illuminated regions unchanged thereby exhibits very limited CC performance as it is shown in (Figure 6-6(l)).

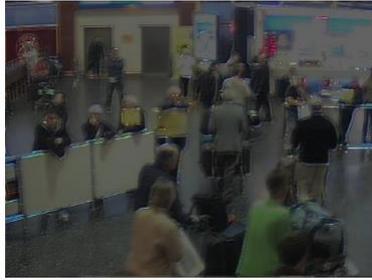
Figure 6-6(m) and Figure 6-6(n) shows the images after processed by Grey-Edge and CBTF methods respectively. The Grey-Edge utilises the global mean of the whole scene to estimate the strength of the light source and thus the method is more applicable to uniformly illuminated scene as it can be seen from Figure 6-6(m). In the CBTF method the camera 1 image has been used as the reference and the brightness is then transferred to camera 3 and camera 5. The histogram of the scene in camera 5 view is quite different from that of camera 1 thereby giving large error in this camera view as shown in Figure 6-6(n).

The ELRCC method previously developed in this work handles illumination pixel-by-pixel but with a drawback of coping high radiance pixels (see Figure 6-6(o)). This problem is overcome by the presently proposed ETDC method which shows better CC over all three different non-overlapping camera views (see Figure 6-6(p) and section 6.4.2 and 6.4.3).





(c)



(d)



(e)



(f)



(g)



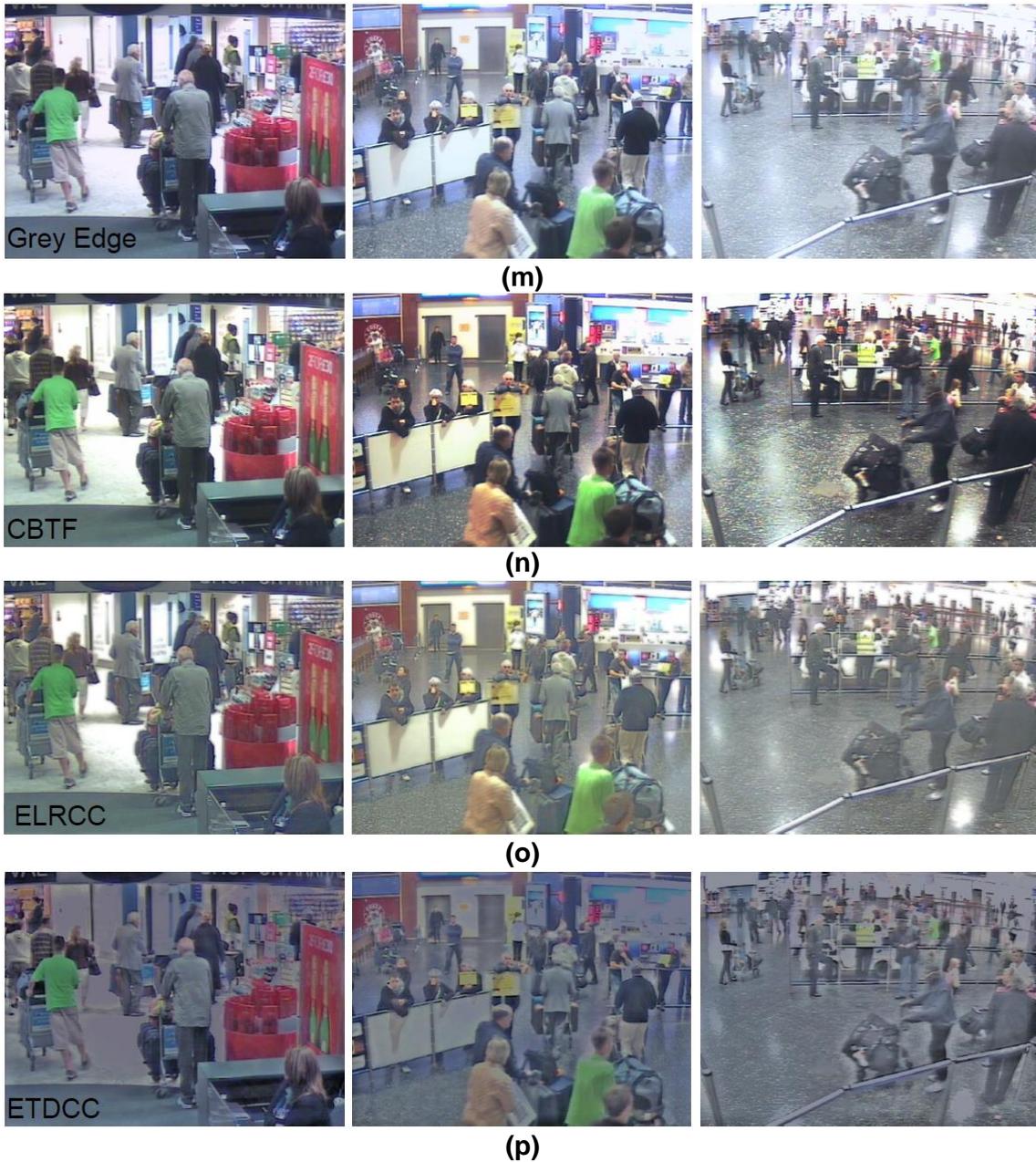


Figure 6-6: Illustrates the RGB image of target 6 for 3 views from iLIDS dataset (a) raw images and (b)-(p) after processed by various CC algorithms.

6.4.2 Colour Descriptors: EDTCC algorithm

Table 6-1 and Table 6-2 tabulate detection results for all targets (16 targets) in the FOM metric using the EDTCC algorithm when it is processed with various colour descriptors. Table 6-3 shows the average data of Table 6-1 and Table 6-2. It is seen that when the EDTCC utilises O1O2O3 colour descriptor it

exhibits the best colour invariance properties with much better detection performance over all other colour descriptors.

Table 6-1 demonstrates the consistence trend of target detection enhancement in all six different descriptors. The sum and L1L2L3 features show considerably poor detection performance with respect to other features which shows that these features are unable to cope with complex illumination variations. Similarly invariant colour features rgb and C1C2C3 also unable to manage colour variations this can be seen when camera 3 data is processed by rgb and C1C2C3 feature, which shows low FOM (i.e. lower detections) as compare to data processed with ETDCC algorithm in camera 3. Moreover due to chrominance conversion the rgb and C1C2C3 descriptors unable to distinguish different shade of grey as discussed above (see section 6.2).

Likewise Table 6-2 tabulates the targets FOM data for six features on PETS dataset which is overlapping camera view. Interestingly the performance of sum and L1L2L3 descriptors has risen noticeably, it is due to fact that the target in overlapping camera view is less challenging as compare to non-overlapping camera view. Again the processed data with proposed method gives the evidence that O1O2O3 descriptor shows significant improvement in target detection that is by 192% when it compares to raw unprocessed data. Over all we can see that ETDCC utilises O1O2O3 colour descriptor exhibits good colour invariance properties with better detection performance the average enhancement rate is found to be 183%

Table 6-1: Tabulates the FOM results for the detection of 12 targets in 3 different non-overlapping camera views of iLIDS data after the images treated by ETDCC+ six different descriptors.

Targets	M:ETDCC	M: Raw	M:ETDCC	M:Raw	M:ETDCC	M:Raw	M:ETDCC	M:Raw	M:ETDCC	M:Raw	M:ETDCC
	F:RGB	F:rgb	F:rgb	F:Sum	F:Sum	F:L123	F:L123	F:C123	F:C123	F:O123	F:O123
T1C1	182.35	192.67	187.80	19.38	47.88	77.44	133.94	126.47	91.38	43.43	211.79
T1C3	223.05	157.05	326.23	109.51	136.93	47.56	256.37	31.46	259.35	23.95	310.83
T1C5	45.31	226.35	244.28	-51.96	-23.60	32.92	183.14	167.57	212.93	20.73	205.84
T2C1	125.65	96.07	97.54	18.04	28.80	32.87	24.54	65.83	55.21	40.56	155.53
T2C3	115.63	73.59	116.52	16.72	46.84	18.79	26.04	48.37	69.45	31.08	186.4
T2C5	383.92	508.14	520.23	111.05	297.61	172.56	150.83	464.69	385.50	62.12	541.94
T3C1	31.55	40.93	40.70	10.33	6.64	14.63	24.05	26.26	20.23	24.76	43.71
T3C3	77.03	62.89	85.17	41.68	65.34	27.03	20.45	33.46	64.32	39.02	105.2
T3C5	50.86	110.52	111.31	-47.15	45.60	-35.94	19.67	94.31	90.34	55.92	110.6
T4C1	72.93	114.09	114.73	46.56	9.95	17.31	28.92	104.44	97.22	28.14	109.11
T4C3	145.03	74.18	194.05	73.01	7.88	39.60	51.10	80.07	192.78	39.33	205.82
T4C5	7.98	63.22	70.23	-56.48	-38.70	-42.10	44.42	49.36	19.80	7.65	113.29
T5C1	54.33	67.45	65.65	4.73	29.03	11.81	20.63	58.63	39.51	21.63	76.46
T5C3	188.84	61.3	229.34	27.58	24.30	124.89	100.53	53.73	181.82	2.37	240.62
T5C5	427.68	601.33	471.54	-78.00	-60.76	340.00	101.72	534.20	479.32	-29.83	543.51
T6C1	49.66	78.73	78.58	19.84	34.28	22.68	20.00	74.34	75.12	98.68	79.18
T6C3	196.17	168.82	192.53	12.79	42.49	15.96	34.07	122.58	182.31	235.86	249
T6C5	-8.26	280.11	288.09	30.80	260.28	20.02	140.77	212.91	270.21	296.76	266.78
T7C1	77.14	73.82	70.18	16.63	93.28	28.01	46.37	49.48	42.11	14.23	92.09
T7C3	125.93	120	138.68	85.89	102.68	9.43	67.89	119.43	126.48	53.62	139.74

Targets	M:ETDCC	M: Raw	M:ETDCC	M:Raw	M:ETDCC	M:Raw	M:ETDCC	M:Raw	M:ETDCC	M:Raw	M:ETDCC
	F:RGB	F:rgb	F:rgb	F:Sum	F:Sum	F:L123	F:L123	F:C123	F:C123	F:O123	F:O123
T7C5	35.14	191.92	181.27	-28.61	147.01	-14.12	71.48	196.20	185.15	6.06	175.89
T8C1	64.6	73.97	75.32	5.31	65.33	2.53	14.77	47.65	37.12	48.37	81.92
Target Absent											
T8C5	34.39	113.08	110.37	-34.33	-30.10	-43.05	6.61	108.08	74.52	75.53	130.61
T9C1	146.64	76.3	75.15	27.88	63.10	43.94	20.40	57.34	52.55	31.45	175.4
T9C3	83.73	39.48	56.50	10.60	56.62	1.11	37.15	25.85	32.97	15.8	111.47
T9C5	241.53	262.55	260.20	1.06	57.95	70.92	42.08	253.11	233.86	28.7	286.46
T10C1	65.43	84.64	82.18	36.76	59.41	22.39	47.00	63.23	55.30	35.78	97.35
Target Absent											
T10C5	-35.51	86.99	86.85	-10.65	-4.84	-31.06	30.52	80.11	74.58	110.37	96.56
T11C1	83.24	80.12	77.15	15.23	62.83	15.78	12.05	56.02	45.02	26.48	107.91
T11C3	72.61	59.54	68.73	13.25	46.60	10.78	26.29	56.43	57.74	19.44	92.89
T11C5	171.35	178.39	170.16	-29.23	58.54	64.90	-0.06	175.92	179.87	31.9	206.64
T12C1	52.08	78.09	79.39	3.19	47.71	4.40	40.56	52.46	43.15	24.11	79.55
Target Absent											
T12C5	104.91	153.31	155.16	-60.60	114.13	-35.97	58.02	151.93	163.81	59.88	144.74
Average	111.90	140.89	155.20	10.93	57.60	32.96	57.64	116.42	127.00	49.20	174.99

Table 6-2: Shows the FOM in % results for the detection of 4 targets in overlapping views of PETS data after images treated by ETDCC+ six different descriptors.

Targets	M:ETDCC	M: Raw	M:ETDCC	M: Raw	M:ETDCC	M: Raw	M:ETDCC	M: Raw	M:ETDCC	M: Raw	M:ETDCC
	F:RGB	F:rgb	F:rgb	F:Sum	F:Sum	F:L123	F:L123	F:C123	F:C123	F:O123	F:O123
T1C1	58.57	56.22	55.64	2.37	48.78	8.63	15.84	57.46	57.41	67.09	77.89
T1C2	31.52	47.71	40.27	23.02	6.49	26.37	26.31	49.90	39.97	79.69	46.3
T1C3	-14.45	84.56	84.70	85.22	13.15	96.29	82.10	80.62	76.83	39.17	77.57
T1C4	164.04	334.07	350.30	-32.80	241.21	144.31	147.02	274.41	347.41	295.86	345.31
T2C1	91.21	114.38	112.29	1.81	62.58	5.77	2.63	133.04	125.68	3.12	165.79
T2C2	70.02	71.38	84.84	90.78	43.27	8.44	4.63	63.84	67.96	23.31	139.59
T3C1	85.74	107.38	107.30	4.38	57.06	11.96	4.45	102.87	102.18	101.12	124.5
T3C2	69.19	79.44	94.63	3.76	4.74	14.38	26.74	32.70	54.34	73.42	177.13
T3C3	-0.16	123.04	139.29	133.39	404.92	700.50	700.75	-13.12	2.73	804.39	786.41
T3C4	108.05	261.38	256.87	-37.17	31.08	43.42	36.08	260.22	253.64	354.07	350.14
T4C1	41.96	69.46	82.93	17.59	43.19	29.16	3.17	69.11	84.13	67.25	83.15
T4C2	37.24	75.32	83.95	1.93	11.87	0.92	1.47	73.05	82.75	104.96	111.41
T4C3	16.84	-88.34	81.92	7.06	51.46	100.64	86.00	24.98	11.29	11.72	45.52
T4C4	57.34	110.9	110.91	45.58	57.82	49.04	41.96	105.43	109.17	77.86	161.98
Average	58.36	103.35	120.41	24.77	76.97	88.55	84.22	93.89	101.10	150.21	192.33

Table 6-3: Tabulates the average FOM in % results for the detection of 16 targets by using ETDCC + various colour descriptors.

Targets	M:ETDCC	M: Raw	M:ETDCC	M: Raw	M:ETDCC	M: Raw	M:ETDCC	M: Raw	M:ETDCC	M: Raw	M:ETDCC
	F:RGB	F:rgb	F:rgb	F:Sum	F:Sum	F:L123	F:L123	F:C123	F:C123	F:O123	F:O123
Average	85.14	122.12	137.81	17.86	67.29	60.76	70.94	105.16	114.05	99.71	183.66

6.4.3 Target detection results by ALL CC algorithms

Table 6-4 and Table 6-5 summarise the detection results for all in the FOM metric by using all CC algorithms utilised in this study. Table 6-6 presents the main result of this project which shows the averaged results of Table 6-4 and Table 6-5. This result has shown an overall superior colour constancy capability of the proposed ETDCC algorithm with respected to all other CC methods utilised in this study.

Table 6-4 shows 12 target enhanced detection data with various CC methods on iLIDS dataset (non-overlapping camera view). According to the data in Table 6-4 all other CC algorithm unable to produce consistent results when compared with raw data. There are considerable difference in the figures of CBTF algorithm and histogram equalisation methods and the FOM in each target in each camera view are in negative values, which indicates that their target detection performance is even poor than raw data. In histogram equalisation methods, histogram stretching in each band causing targets colours to washed out and while in CBTF method the histogram of camera 1 is transferred to camera 5 however local (each pixel) information does not matched. The LBMSR method in Table 6-4 shows better target detection result as compare to Retinex based method because LBMSR technique is based on luminance of an image instead of multiple surround function on each colour attributes of the image as in the Retinex methods. Table 6-4 also shows that NLTF based methods have low FOM figures (i.e. lower detection) specially in IRN and NWIS methods which exhibits poor target detections, this is mainly due to the use of global parameters (i.e. global mean I_m of the whole scene). The FOM highlights that these techniques unable to cope with colour variations and hence fails the target detection in 3 difference camera view with complex illumination.

The ELRCC framework as defined in chapter 4 out-performs all other methods which are not proposed in this research work and shows considerable enhancement in target detection with respect to all Retinex and CC (white balancing) techniques because of pixel-wise methodology. The ETDCC figures shows that the detection performance has risen significantly up to 175% and

shows better detection results than ELRCC, subsequently with all other CC algorithms. Moreover if we compare the pixel-based NLTF with fixed parameter approach that is fixed alpha, then there are noticeable differences between these figures and it highlights that the pixel-wise approach is far more robust and reliable. Likewise if we compare OneLog enhancement approach that is one directional NLTF curves with ETDC (bidirectional NLTF), which handling both low and high intensities, then the FOM in Table 6-4 suggest that bidirectional approach is critically important and shows substantial target detection results that is 175% enhancement as compare to 100% by using OneLog function.

Similarly Table 6-5 shows targets enhanced detection FOM data results for various other CC methods on PETS dataset (overlapping camera view). Table 6-5 follows the same pattern as Table 6-4 and ELRCC out performed all other CC algorithm which are not proposed in this research work. Afterwards ETDC shows marginally better detection results as compare to ELRCC. However the performance of all other methods, especially CBTF, Histogram Equalisation, Retinex and CC (white balancing methods) are rose dramatically. This is due to the main reason of overlapping camera view which technically specking pose less challenging as compare to non-overlapping camera view where, illumination, scene, background and conditions are totally different. Table 6-6 shows the average data Table 6-4 and Table 6-5 that is iLIDS and PETS datasets respectively. An overall summary of the performance of 15 different methods in two different datasets shows the evidence that ETDC exhibits considerably good colour invariant characteristic and therefore demonstrates better target detection in overlapping and non-overlapping camera view as compare to other CC methods.

Table 6-4: Compares the FOM in % for the detection of targets in the non-overlapping camera iLIDS data set after images processed by various CC algorithms.

Target	ETDCC	ELRCC	Grey World	Grey Edge	MSRCR	LBMSR	IRN	AIND-ANE	NWIS	Fix Alpha	One Log	CBTF	Hist EQ C	Hist EQ Int	Avg Int
T1C1	211.79	108.98	44.85	38.13	129.31	129.91	69.56	69.83	90.51	93.57	119.44	40.69	78.88	32.22	39.72
T1C3	310.83	123.25	23.05	87.52	230.89	165.91	71.33	66.71	29.34	76.68	123.19	-20.53	-4.88	40.60	106.51
T1C5	205.84	82.59	-7.29	53.20	-88.80	-33.85	50.01	62.64	34.79	106.14	132.18	-52.23	-0.46	27.21	141.99
T2C1	155.53	62.47	42.82	32.51	21.54	59.03	50.07	29.62	73.89	37.68	54.19	45.24	69.48	8.46	7.64
T2C3	186.40	65.49	50.52	8.51	66.21	21.73	60.67	72.87	28.40	-6.15	4.10	54.52	55.46	28.72	63.86
T2C5	541.94	296.94	94.47	-40.07	455.41	354.62	206.45	232.57	134.98	189.50	279.76	-65.65	-7.68	1.63	378.53
T3C1	43.71	37.17	27.02	20.66	57.28	44.41	24.17	14.23	34.25	26.03	36.09	23.88	34.41	24.81	19.79
T3C3	105.20	50.55	34.54	67.82	84.11	66.31	40.85	-4.62	32.75	22.84	40.75	-13.63	-4.24	10.29	70.17
T3C5	110.60	78.66	46.51	65.89	2.67	78.29	62.17	24.95	73.76	71.75	94.11	-33.21	-15.95	36.49	85.88
T4C1	109.11	91.20	17.20	29.73	46.09	52.02	48.30	83.46	40.32	42.34	50.02	19.06	1.99	33.82	29.16
T4C3	205.82	178.52	67.54	25.76	114.22	115.64	121.09	192.99	34.75	83.96	100.89	177.09	119.05	97.31	205.21
T4C5	113.29	126.28	28.56	-2.90	92.04	88.96	90.61	119.29	28.73	107.78	103.34	22.56	15.43	65.47	131.74
T5C1	76.46	46.35	21.59	29.82	52.66	66.67	36.44	32.62	43.47	25.75	45.12	21.56	58.00	21.77	5.72
T5C3	240.62	52.02	1.87	97.62	115.72	101.73	32.17	32.06	3.35	20.55	64.37	-17.57	14.97	48.33	62.51
T5C5	543.51	110.63	-42.72	30.85	-68.29	14.72	51.91	49.31	-19.32	52.26	145.49	-75.01	-58.29	128.55	190.24
T6C1	79.18	79.39	98.35	99.75	94.88	85.22	92.95	90.23	86.73	96.54	95.34	96.98	75.15	91.98	72.40
T6C3	249.00	223.17	243.39	256.25	218.87	207.06	240.85	234.52	234.01	238.17	226.36	195.01	190.82	232.45	147.54
T6C5	266.78	309.50	276.77	311.47	293.15	256.38	307.47	314.47	298.85	302.07	246.75	207.49	210.11	347.00	237.86
T7C1	92.09	62.38	20.62	19.18	58.22	67.74	41.29	68.18	33.37	37.80	50.35	18.35	55.25	24.84	7.81
T7C3	139.74	109.51	50.14	89.34	162.23	140.47	87.59	106.99	59.26	75.90	105.20	105.03	116.44	55.44	93.44
T7C5	175.89	129.04	-35.83	37.68	83.21	144.71	78.42	155.82	74.48	79.17	113.27	-70.59	-31.99	80.64	131.28

Target	ETDCC	ELRCC	Grey World	Grey Edge	MSRCR	LBMSR	IRN	AIND-ANE	NWIS	Fix Alpha	One Log	CBTF	Hist EQ C	Hist EQ Int	Avg Int
T8C1	81.92	69.29	41.73	53.74	67.52	75.15	63.94	64.16	72.62	78.39	84.64	45.31	65.48	51.18	38.25
T8C3	Target Absent														
T8C5	130.61	108.51	55.75	95.37	-20.02	84.56	106.24	130.78	90.04	120.88	117.84	-56.43	-45.46	112.19	119.30
T9C1	175.40	82.95	32.80	9.33	31.22	86.00	66.92	86.47	72.47	61.74	75.30	20.53	16.57	64.96	12.65
T9C3	111.47	45.73	17.90	-9.23	42.94	20.65	41.95	72.33	-2.43	-12.97	-6.38	6.72	-13.89	24.83	51.56
T9C5	286.46	213.45	32.33	-21.49	138.59	182.16	130.69	212.15	68.42	164.89	177.18	-62.15	-63.41	126.57	259.70
T10C1	97.35	30.75	34.16	37.67	64.74	73.50	34.46	42.19	66.55	75.72	87.27	28.50	48.78	56.13	39.72
T10C3	Target Absent														
T10C5	96.56	118.00	106.72	108.15	34.97	69.39	115.19	123.26	97.37	122.03	86.96	-52.23	-37.26	136.12	92.18
T11C1	107.91	91.56	31.53	8.27	51.83	69.07	51.33	103.20	57.15	65.86	72.40	38.64	14.55	10.12	9.31
T11C3	92.89	74.83	5.36	-26.24	55.67	32.16	46.41	107.50	13.86	24.01	30.21	7.95	-5.53	6.03	56.30
T11C5	206.64	199.80	22.93	-16.95	75.97	144.63	126.84	286.76	102.51	158.81	159.33	-57.24	-66.76	32.35	254.97
T12C1	79.55	43.52	25.10	28.78	90.06	53.51	31.24	53.06	43.06	49.69	59.41	26.64	30.83	50.60	24.58
T12C3	Target Absent														
T12C5	144.74	89.64	41.73	93.54	125.46	111.30	76.55	145.58	94.97	140.28	129.55	-89.25	-86.37	140.81	159.25
Avg	174.99	108.85	47.03	52.11	90.32	97.87	83.52	105.34	67.49	85.75	100.12	16.24	25.14	68.18	101.42

Table 6-5: Compares the FOM in % for the detection of targets in the overlapping camera PETS data set after images processed by various CC algorithms.

Target	ETDCC	ELRCC	Grey World	Grey Edge	MSRCR	LBMSR	IRN	AIND-ANE	NWIS	Fix Alpha	One Log	CBTF	Hist EQ C	Hist EQ Int	Avg Int
T1C1	77.89	83.90	66.52	66.81	81.00	82.48	80.92	66.52	65.78	86.10	83.17	69.27	70.95	73.42	65.15
T1C2	46.30	54.97	70.18	72.16	53.87	71.51	75.34	55.41	65.34	81.48	80.07	75.89	74.97	77.76	88.08
T1C3	77.57	76.53	46.04	61.34	89.34	80.08	77.53	50.79	58.00	72.96	75.19	-44.33	-32.07	65.77	32.55
T1C4	345.31	384.55	314.52	311.69	374.02	404.79	401.82	404.70	332.55	411.89	408.35	251.60	276.73	297.45	258.19
T2C1	165.79	86.13	7.74	7.56	52.14	65.52	79.35	178.69	21.60	108.27	107.70	2.00	34.56	72.85	40.44
T2C2	139.59	101.97	34.62	14.83	11.16	85.75	93.27	112.41	45.36	101.08	93.99	60.43	21.21	103.48	9.74
T3C1	124.50	123.41	104.68	103.47	103.72	113.64	122.33	93.64	109.18	120.12	120.32	101.89	89.42	97.80	65.15
T3C2	177.13	177.00	131.78	123.48	163.95	179.02	131.13	134.53	134.33	149.71	151.62	57.06	94.86	142.97	88.08
T3C3	786.41	762.30	650.13	737.34	554.17	656.01	768.82	696.96	683.62	848.97	823.93	401.11	570.41	653.12	341.79
T3C4	350.14	390.57	327.82	346.42	349.46	333.38	362.50	345.00	372.37	379.03	371.61	259.75	244.44	312.47	258.19
T4C1	83.15	86.29	75.63	73.03	74.06	88.27	74.82	46.21	62.13	76.52	83.43	68.46	47.54	26.98	37.50
T4C2	111.41	95.09	115.53	108.58	107.66	115.35	104.85	90.09	79.94	109.65	111.82	81.87	48.55	72.47	60.12
T4C3	45.52	18.40	67.35	33.64	-71.24	51.34	36.33	16.61	-49.53	-68.99	-50.37	16.14	46.00	26.35	7.01
T4C4	161.98	129.65	68.86	82.26	128.24	145.25	152.01	143.00	98.46	153.07	150.61	128.05	83.59	165.58	109.80
Avg	192.34	183.63	148.67	153.04	147.97	176.60	182.93	173.90	148.51	187.85	186.53	109.23	119.37	156.32	104.41

Table 6-6: Shows the main result of this work which tabulates the FOM in % for the detection of targets in the iLIDS and PETS data set after images processed by various CC algorithms.

Target	ETDCC	ELRCC	Grey World	Grey Edge	MSRCR	LBMSR	IRN	AIND-ANE	NWIS	Fix Alpha	One Log	CBTF	Hist EQ C	Hist EQ Int	Avg Int
Avg	177.15	142.05	93.40	98.30	115.75	131.43	126.29	135.85	104.63	130.83	137.51	59.64	68.13	112.24	102.91

6.5 GTCC data set: All CC algorithms

6.5.1 Colour enhancements

Figure 6-7 highlights the ability of all CC algorithms studied in this work using 4 samples of images extracted from the GTCC data set. The sample images were taken under 4 different irradiances as described in section 3.1.3 (left to right, low to high irradiance as shown in raw images in Figure 6-7(a)).

Figure 6-7(b)-(e) outline the images after processed by Histogram Equalisation, MSR, MSRCR and LBMSR respectively. All these results have exhibited substantial colour bleach unable to maintain colour constancy although MSRCR and LBMSR have improved the colour contrast somewhat through the centre and surround operation. Parameterisation in Retinex is normally obtained by experimental trials and inappropriate parameters will cause bad results as shown in Figure 6-6(f) and Figure 6-6(g). The image after processed by NWIS exhibits very dark background (Figure 6-7(h)) due to the excessively mid-tone compression.

The results by ETDCC method depicts much better CC results after the images pass the 3 stages of processing: PWNLTF, CSL and grey edge method (Figure 6-7(l)). It is seen that processing by a single method such as the Grey-Edge method as shown in Figure 6-7(m) is not adequate enough to maintain CC.

Figure 6-8 shows the mean RGB of each frame of image taken under 10 different irradiances for the experiment as described in section 3.1.3. Note that the scene is the same and only the illumination is different, thus it is expected that a good CC algorithm will be able to maintain the mean intensity of the scene (i.e. mean [RGB] of each frame) to be a constant. It is quite clear that the proposed ETDCC algorithm is capable to maintain colour constancy rather robustly.

AINDANE



(g)

NWIS



(h)

Fix Alpha



(i)

One Log



(j)

ELRCC



(k)

ETDCC



(l)

Grey-Edge



Figure 6-7: Illustrates the RGB image of 4 images from the GTCC dataset (a) raw images and (b)-(m) after processed by various CC algorithms.

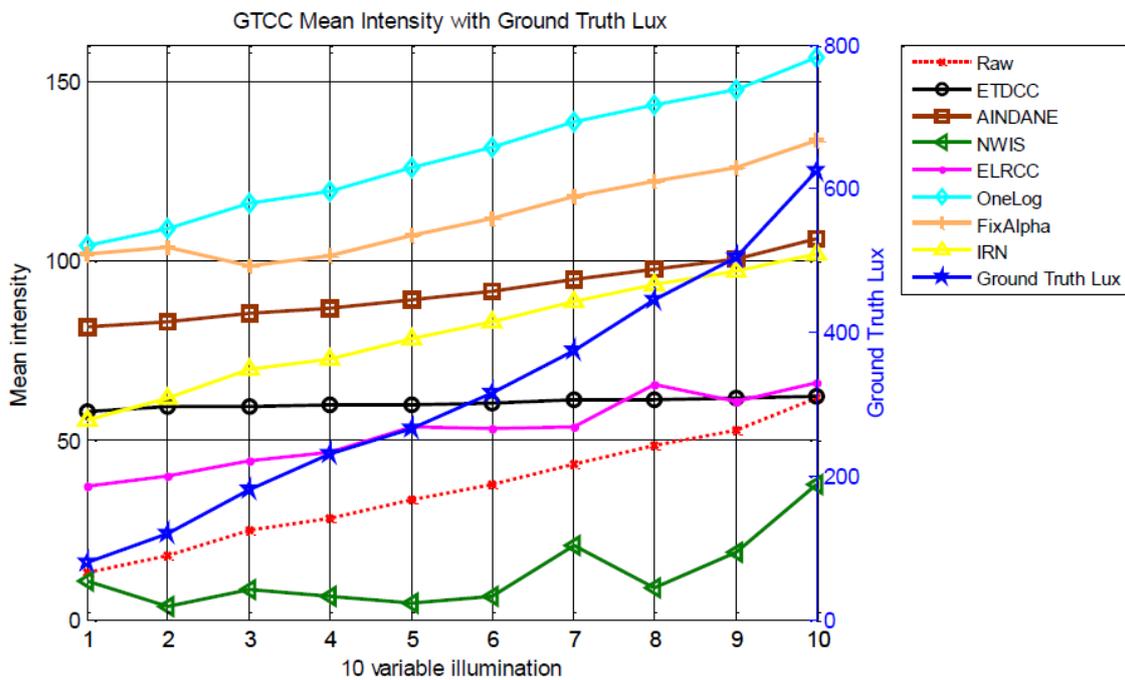


Figure 6-8: Shows the mean intensity of each frame of image extracted from the GTCC data set after they are processed by various CC algorithms. The x-axis is the 10 experimental runs for recording the scene under 10 different irradiances.

6.5.2 GTCC data set: ΔE_{ab} assessment

Table 6-7 tabulates the ΔE_{ab} of 24 colour boxes of the GTCC dataset and the colour boxes have been used as targets in this experiment. As described in section 3.5.2 the ΔE_{ab} metric has been utilised for assessing the colour or chrominance difference of the image frame with respected to the ground truth.

Each colour box has 10 ΔE_{ab} values and each is evaluated from the image taken at step change of irradiance as detailed in section 3.1.3. In each entry of the ΔE_{ab} as shown in the table, it shows the average of these 10 ΔE_{ab} values as function of various CC algorithms utilised for treating the data. The average of the ΔE_{ab} for all 24 targets is tabulated in the row labelled averaged $\sum \Delta E_{ab}$ (highlighted in green) in Table 6-7. Note that the smaller the ΔE_{ab} the better the colour constancy ability of the algorithm. In Table 6-7 there is noticeable difference in the figures of Grey-Edge, Grey-World and NWIS methods which shows quite the same ΔE_{ab} values as raw data. The ELRCC method shows better performance than raw and CC methods, but quite the opposite, Retinex and NLTF methods out performs ELRCC method in GTCC dataset. However the proposed ETDCC method shows significantly better colour difference results as compare to all other CC methods by average of 24 colour boxes $\sum \Delta E_{ab}$ is 173, which indicates that the ETDCC minimise the illumination artifacts in 10 variable illumination images, thus keeping the colour difference at lowest level and improved the similarity in variable illumination images, consequently better CC in these images. It is observed that the least ΔE_{ab} is found from the data after it is treated by the HistEQ algorithm: it shows almost 4 times less than that of the proposed ETDCC method which indicates the HistEQ is the best colour constancy algorithm. It can be seen from Figure 6-7 that although the Histogram Equalisation method has shown rather good colour constancy (i.e. very low ΔE_{ab}), the colours of the images after processed by HistEQ are found strongly bleached. Table 6-6 shows very clear that the target detection capability of HistEQ is about halve of that by ETDCC.

Table 6-7: Tabulates the ΔE_{ab} for processing the images in the GTCC data set by various CC algorithms. Each entry is the averaged ΔE_{ab} colour differences of 10 images recorded under step change of irradiance.

Colour Box	Raw	ETDCC	ELRCC	Grey World	Grey Edge	MSR	MSRCR	LBMSR	IRN	AINDANE	NWIS	Fix Alpha	One-Log	Hist EQ
	1633	71	495	1624	1620	408	384	376	579	368	3125	397	423	53
	1352	225	553	1333	1345	223	202	267	559	187	2704	383	402	2
	1515	58	317	1472	1507	282	235	195	383	71	2896	193	210	13
	1626	88	388	1592	1613	379	335	284	450	158	3099	250	281	39
	1430	85	317	1395	1421	245	206	188	374	54	2787	176	190	55
	1283	86	341	1227	1262	190	130	135	364	155	2618	164	175	54
	1495	503	771	1503	1492	383	480	456	747	357	2857	570	579	82
	1553	158	400	1511	1545	327	297	272	496	165	3008	293	306	37
	1531	406	706	1536	1542	363	397	497	717	238	2936	523	537	56
	1639	88	525	1625	1645	428	384	329	656	758	3184	473	503	61
	1364	235	482	1327	1350	239	224	255	521	337	2716	337	351	2
	1394	430	715	1392	1393	320	393	361	692	398	2750	511	521	78
	1639	109	412	1588	1617	422	390	303	531	897	3190	323	334	70
	1508	91	387	1457	1492	299	241	258	469	127	2895	287	305	10
	1624	398	691	1666	1629	530	609	513	744	388	3069	555	573	203
	1261	352	641	1239	1261	239	281	284	615	432	2653	431	440	33
	1444	287	587	1452	1450	318	324	377	606	132	2827	417	429	41
	1475	139	360	1413	1452	286	229	212	417	55	2865	219	229	42
	881	44	381	844	872	78	29	46	307	294	2329	133	142	8
	1092	14	326	1054	1082	133	81	89	331	302	2472	144	159	0.5
	1299	54	299	1265	1290	196	145	138	341	133	2654	155	175	0
	1503	69	350	1462	1495	285	233	211	412	72	2896	222	246	0.2
	1614	66	401	1584	1616	383	335	297	462	277	3152	279	307	0.24
	1714	97	649	1677	1714	583	550	543	760	514	3277	539	586	169
Average	1453	173	479	1427	1446	314	296	287	522	286	2873	332	350	46

6.5.3 GTCC data set: ETDCC vs colour descriptors

This section presents results similar to that in Table 6-1 and Table 6-2 but using the GTCC data instead. The CDP assessment method has been adopted in this test. CDP compares the colour difference similar to that of ΔE_{ab} but it is in percentage metric (see section 3.5.3).

Table 6-8 tabulates the CDP of 18 different colour boxes of the GTCC dataset as targets. Each colour box has 10 CDP values and each is evaluated from the image taken at step change of irradiance as detailed in section 3.1.3. In each entry of the CDP as shown in the table, it shows the average of these 10 CDP values as function of various CC algorithms utilised for treating the data. The average of the CDP for all 18 targets is tabulated in the row labelled averaged (highlighted in green) in Table 6-8. Note that the smaller the CDP the better the colour constancy ability of the algorithm.

Table 6-8 shows the CC ability of the ETDCC with six different colour descriptors using the GTCC data set. It can be seen from the table that the best colour constancy is in fact achieved by the ETDCC when the reduced rgb colour descriptor is employed: it shows the least CDP of 1.75. Despite of having very high CC ability in the ETDCC when the reduced rgb descriptor is employed, it suffers greatly from colour bleaching and exhibits low target detection performance as evidence from Table 6-3 which shows ~50% lower of the target detection in reduced rgb than that using O1O2O3 descriptor.

Table 6-8: Tabulates the CDP for processing the images in the GTCC data set by the ETDC algorithm together with various colour descriptors. Each entry is the averaged CDP over 10 images recorded under step change of irradiances.

Colour Box	Raw F:RGB	M:ETDCC F:RGB	M:Raw F:O123	M:ETDCC F:O123	M:ELRCC F:O123	M:Raw F:C123	M:ETDCC F:C123	M:Raw F:rgb	M:ETDCC F:rgb	M:Raw F:Sum	M:ETDCC F:Sum	M:Raw F:L123	M:ETDCC F:L123
	8.64	0.96	7.52	2.23	2.55	1.74	1.23	1.37	0.98	23.15	16.84	3.29	3.92
	7.56	2.05	5.22	4.57	3.65	2.44	1.96	1.85	1.45	1.88	3.63	2.35	3.49
	8.18	0.59	8.56	1.20	3.13	0.51	0.53	0.48	0.45	4.73	6.74	2.65	2.53
	8.58	0.88	8.61	1.20	3.30	0.68	0.39	0.59	0.37	35.04	26.75	3.63	1.94
	7.85	0.83	7.02	4.17	5.68	0.69	0.64	0.81	0.63	2.98	4.07	16.68	6.30
	7.55	1.61	7.35	1.18	2.78	1.04	1.26	0.88	1.15	4.01	4.73	20.70	12.26
	8.40	4.57	6.06	1.99	2.74	4.65	4.27	3.96	3.62	18.27	27.03	4.15	3.68
	8.33	1.47	9.50	3.14	5.74	0.98	1.36	0.84	1.03	6.99	10.02	5.57	4.65
	8.47	3.86	6.22	2.17	2.58	4.07	3.66	3.23	2.87	12.68	17.82	5.43	4.02
	8.68	1.03	9.22	2.32	5.22	1.30	0.75	0.98	0.65	19.46	15.13	623.41	198.50
	7.89	2.57	6.20	3.48	2.49	2.11	1.80	2.30	2.08	9.78	13.74	2.62	0.76
	8.04	4.13	5.59	3.12	3.27	4.27	3.88	3.69	3.36	9.94	15.71	6.57	5.15
	8.66	1.14	8.90	1.55	3.94	0.79	1.19	0.66	0.93	14.99	15.30	2.01	1.73
	8.42	1.49	6.98	6.27	5.12	1.11	1.23	1.35	1.23	42.91	47.51	4.53	3.48
	9.07	4.42	6.94	1.28	2.39	4.78	4.41	4.14	3.80	30.51	39.88	6.27	4.60
	7.47	3.46	5.41	4.42	4.63	3.56	3.17	3.14	2.83	8.29	12.30	11.67	6.96
	8.14	3.11	6.82	8.87	8.68	3.35	2.91	2.63	2.31	4.48	7.38	3.99	2.79
	8.31	2.28	7.89	0.95	2.42	1.56	2.06	1.37	1.78	6.54	9.63	2.56	3.78
Avg	8.24	2.25	7.22	3.01	3.91	2.20	2.04	1.90	1.75	14.26	16.35	40.45	15.03

Chapter 7: Conclusion and Future Works

This research concerns with the development of an effective colour constancy (CC) algorithm to enhance detection of targets in multi-view non-overlapping camera network system which uses spectral characteristics of targets as the means of detection. The programme is aimed to answer the following questions:

1. To understand whether target detection in multi-view camera system can be benefited from CC technique?
2. To understand the drawbacks of existing CC algorithms why they cannot be utilised for practical CCTV multi-camera network system?
3. How to develop a CC algorithm which is capable to maintain the spectral (colour) of the images to a constant irrespective of illumination and camera calibration issues, and at the same time to enhance target detection?

To answer the above questions two different forms of CC algorithms have been developed in this programme and their performances together with 11 other CC algorithms reported in the literature, have been critically tested via 3 different data sets to verify their CC capabilities as well as their target detection efficiencies. A total of 40 targets selected from 3 data sets of a. non-overlapping camera view real iLIDS dataset, b. overlapping camera view PETS 2007 dataset & c. GTCC dataset, have been utilised for assessing the CC and target detection abilities by three metrics of: i) figure of merit (FOM) based upon Area Under Receiver Operating Characteristics (AUROC), ii) Colour differences ΔE_{ab} and iii) Colour Difference Percentage (CDP).

The main result of this research can be summarised as follows:

1. Most CC algorithms, such as the Retinex and adaptive histogram equalisation (HistEQ), have been shown to possess high colour constancy capabilities but they suffer from drawbacks of having strong colour bleaching side effects. The heavily colour bleach makes target detection difficult especially when colour is used as the detection feature. The cause of the colour bleach is mainly due to the lack of a principle

way for the parameterisation of these algorithms especially in the Retinex based CC algorithms.

2. The use of colour invariant descriptors alone is not sufficient enough to maintain CC and most of them are found effective only to strong colours and ineffective to shades of grey.
3. This research has developed two CC algorithms and the first one is termed as Enhanced Luminance Reflectance Colour Constancy (ELRCC) method which incorporates an effective pixel-wise mid-tone dynamic range compression based on a modified sigmoid function. The pixel-wise mechanism has been shown effective for maintaining CC even in situations of strong local illumination variations. As the result the ELRCC is shown to enhance target detections by 100% better than that using the untreated raw data. A summary of this work has been published in the Optical Engineering journal.
4. The second CC algorithm developed in this research is the Enhanced Target Detection by using Colour Constancy (ETDCC) algorithm which incorporates a bidirectional pixel-wise non-linear transfer function for mid-tone dynamic compression (PWNLTF), a centre surround luminance (CSL), a colour invariant descriptor ($O_1 O_2 O_3$) and a grey edge colour correction. This methodology has been shown capable to maintain colour constancy 1) even in situations that suffer from strong local non-uniform illuminations 2) in multi-camera network with unknown camera calibrations via brightness transfer between different camera views without prior information. The ETDCC algorithm has been shown to enhance target detections by 175% better than that using the untreated raw data, and with detection efficiency far better than all other 11 CC algorithms reported in the literature. A summary of this work is being written up for a journal paper publication.

7.1 Future works

Future work in this direction of research has been identified as follows:

1. Targets with weak colours like grey/white/black have been found difficult to detect using spectral characteristic as a means of detection. Other features such as texture, shape/ appearance and perhaps position base models will be needed for robust detections of these difficult targets. These additional features can help discriminate similar coloured targets present in the scene.
2. All detection performed in this research has been using a single colour as signature per target. Multiple colour tagging such as spectral characteristics of the top/shirt, trousers, hand bags/luggage etc. will help target discrimination immensely.
3. Background segmentation: all detection carried out in this work was performed without background segmentation and it is expected that the false positive can be substantially reduced via background characterisation technique. This methodology may be particularly relevant for the detection of weak coloured targets.
4. Extensions of the ETDCC for across camera synchronisations within and or between CCTV network systems. This can be done by improving the iteration process of Beta, which gives much better control in a low quality video footage. However the computation cost of the process is high approximately 1.3 fps. However one should attempt to employ a training base beta computation process which will reduce the computation cost.

REFERENCES

- Ahlberg, J. and Renhorn, I., 2004. *Multi and hyperspectral target and anomaly detection*. Linkoping, Sweden, Swedish Defence Research Agency.
- Andrew, G. and Bowden, R., 2006. *Tracking objects across cameras by incrementally learning inter-camera colour calibration and patterns of activity*. Graz, Austria, European Conference on Computer Vision (ECCV).
- Arigela, S. and Asari, K., 2006. *An Adaptive and Non Linear Technique for Enhancement of Extremely High Contrast Images*. Washington, USA, IEEE Applied Imagery and Pattern Recognition Workshop.
- Arigela, S. and Asari, V. K., 2013. Self-tunable transformation function for enhancement of high contrast color images. *SPIE Journal of Electronic Imaging*, 22(2), pp. 023010-1- 023010-22.
- Asari, K. V., Oguslu, E. and Arigela, S., 2006. Nonlinear Enhancement of Extremely High Contrast Images for Visibility Improvement. In: P. K. Kalra and S. Peleg, eds. *Computer Vision, Graphics and Image Processing*. Berlin, Heidelberg: Springer, pp. 240-251.
- Bak, S., Corvee, E., Bremond, F. and Thonnat, M., 2010. *Person Re-identification Using Spatial Covariance Regions of Human Body Parts*. Boston, USA, IEEE advanced Video and Signal Based Surveillance (AVSS).
- Barnard, K., 1997. *Computational Color Constancy: Taking Theory into Practice*, British Columbia, Canada: Simon Fraser University , School of Computing Science.
- Barnard, K. and Funt, a. B., 1997. *Analyzing and improving multi-scale retinex*. Arizona, ISandT/SID Fifth Color Imaging Conference: Color Science, Systems and Applications.
- BBC News, 2006. *news.bbc.co.uk*. [Online] Available at: <http://news.bbc.co.uk/1/hi/uk/6108496.stm> [Accessed 2012].

Buchsbaum, G., 1980. A spatial processor model for object colour perception. *Journal of the Franklin Institute*, 310(1), pp. 1-26.

Cardei, V., Funt, B. and Barnard, K., 2002. Estimating the Scene Illumination Chromaticity Using a Neural Network. *Journal of the Optical Society of America*, 19(12), p. 2374–2386.

CCTV, 2014. *CCTV Remote My Home*. [Online] Available at: <http://www.rmhgroup.co.uk/products/security-2/cctv/> [Accessed 1 May 2014].

Chao, W., Cho, CW, Shih, YY, Chen, YY and Chang, C, 2007. Correction of Inhomogeneous MR Images Using Multiscale. *International Journal of Image Processing*, 1(1), pp. 1-16.

Chen, K., Lai, C., Hung, Y. and Chen, C., 2008. *An adaptive learning method for target tracking across multiple cameras*. Anchorage, Alaska, USA, IEEE Computer Vision and Pattern Recognition (CVPR).

Choudhury, A. and Medioni, G., 2009. *Perceptually Motivated Automatic Color Contrast Enhancement*. Kyoto, IEEE International Conference on Computer Vision Workshops.

Choudhury, A. and Medioni, G., 2010. Perceptually Motivated Automatic Color Contrast Enhancement Based on Color Constancy Estimation. *EURASIP Journal on Image and Video Processing*, 2010(837237), pp. 1-22.

Ciurea, F. and Funt, a. B., 2003. *A Large Image Database for Color Constancy Research*. Scottsdale, Proceedings of the Eleventh Color Imaging Conference.

Ciurea, F. and Funt, B., 2004. Tuning Retinex parameters. *SPIE Journal of Electronic Imaging*, 13(1), pp. 58-64.

Dempster, A. P., Laird, N. M. and Rubin, D. B., 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39(1), pp. 1-38.

Eisenbach, M., Kolarow, A., Schenk, K., Debes, K. and Gross, H, 2012. *View Invariant Appearance-Based Person Reidentification Using Fast Online Feature Selection and Score Level Fusion*. Beijing, IEEE Ninth International Conference on Advanced Video and Signal-Based Surveillance (AVSS).

Ess, A., B-Leibe and Gool, L. V., 2007. *Depth and appearance for mobile scene analysis*. Rio de Janeiro, Brazil, IEEE International Conference on Computer Vision.

Farenzena, M., Bazzani, L., Perina, A., Murino, V and Cristani, M, 2010. *Person re-identification by symmetry-driven accumulation of local features*. San Francisco, USA, IEEE Computer Vision and Pattern Recognition (CVPR).

Finlayson, G. H., 1996. Color in Perspective. *IEEE Transactions on Pattern Analysis and Machine*, 18(10), pp. 1034-1038.

Finlayson, G. and Schaefer, G., 2001a. *Convex and non-convex illumination constraints*. Norwich, IEEE Conference on Computer Vision and Pattern Recognition.

Finlayson, G. and Schaefer, G., 2001b. Solving for colour constancy using a constrained. *International Journal of Computer Vision*, 42(3), pp. 127-144.

Finlayson, G. and Trezzi, a. E., 2004. *Shades of gray and colour constancy*. s.l., Proc. ISandT/SID's Color Imag. Conf..

Forsyth, D. A., 1990. A novel algorithm for color constancy. *International Journal of Computer Vision*, 5(1), pp. 5-35.

Freeman, W. T. and Brainard, D. H., 1995. *Bayesian Decision Theory, the Maximum Local Mass Estimate, and Color Constancy*. Massachusetts, USA, IEEE 5th International Conference on Computer Vision.

Funt, B. and Finlayson, G., 1995. Color constant color indexing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(5), pp. 522-529.

Funt, B. V. and Cardei, V., 1999. *Bootstrapping Color Constancy*. San Jose, SPIE Human Vision and Electronic Imaging IV.

- Gehler, P., Rother, C. and Blake, A., 2008. *Bayesian color constancy revisited*. Anchorage, IEEE Computer Vision and Pattern Recognition (CVPR).
- Gevers, T. and Smeulders, A. W. M., 1999. Color-based object recognition. *Springer Pattern Recognition*, 32(3), pp. 453-464.
- Gheissari, N., Sebastian, T. and Hartley, R., 2006. *Person reidentification using spatiotemporal appearance..* New York, USA, IEEE Computer Vision and Pattern Recognition (CVPR).
- Ghimire, D. and Lee, J., 2011. Nonlinear transfer function-based local approach for color image enhancement. *IEEE Transactions on Consumer Electronics*, 57(2), pp. 858-865.
- Ghimire, D. and Lee, J., 2012. Nonlinear Transfer Function-Based Image Detail Preserving Dynamic Range Compression for Color Image Enhancement. In: Y. Ho, ed. *Advances in Image and Video Technology*. Berlin Heidelberg: Springer, pp. 1-12.
- Gijssenij, A. and Gevers, T., 2011. Color Constancy Using Natural Image Statistics and Scene Semantics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(4), pp. 687-698.
- Gijssenij, A., Gevers, T. and Weijer, J. v. d., 2010. Generalized Gamut Mapping using Image Derivative Structures for Color Constancy. *International Journal of Computer Vision*, 86(2-3), pp. 127-139.
- Gijssenij, A., Rui, L. and Gevers, a. T., 2012. Color constancy for multiple light sources. *IEEE Trans. Image Process*, 21(2), p. 697–707.
- Gray, D., Brennan, S. and Tao, H., 2007. *Evaluating appearance models for recognition, reacquisition, and tracking*. Rio de Janeiro, IEEE International Workshop on Performance Evaluation for Tracking and Surveillance.
- Gray, D. and Tao, H., 2008. *Viewpoint Invariant Pedestrian Recognition with an Ensemble of Localized Features*. Berlin, Heidelberg, 10th European Conference on Computer Vision.

- Hamdoun, O., Moutarde, F., Stanciulescu, B. and Steux, B., 2008. *Interest points harvesting in video sequences for efficient person identification*. Marseille, The Eighth International Workshop on Visual Surveillance.
- Hordley, S. and Finlayson, G., 2006. Reevaluation of color constancy algorithm performance. *Journal of The Optical Society of America*, 23(1), pp. 1008-1020.
- Hurlbert, A. C., 1989. *The Computation of Color*, s.l.: Massachusetts Institute of Technology.
- Jang, J. H., Choi, B., Kim, S. D. and Ra, J. B., 2008. *Sub-band decomposed multiscale retinex with space varying gain*. San Diego, USA, IEEE International Conference on Image Processing.
- Javed, O., Shafique, K., Rasheed, Z. and Shah, M., 2008. Modeling inter-camera space time and appearance relationships for tracking across non-overlapping views. *Computer Vision and Image Understanding*, 109(2), pp. 146-162.
- Javed, O., Shafique, K. and Shah, M., 2005. *Appearance modeling for tracking in multiple non-overlapping cameras*. San Diego, IEEE Computer Vision and Pattern Recognition (CVPR).
- Jobson, D. J., Rahman, Z.-u. and Woodell, G. A., 1996. *Retinex Image Processing: Improved Fidelity to Direct Visual Observation*. s.l., Color Imaging Conf. Color Sci. Syst. Appl.
- Jobson, D. J., Rahman, Z.-u. and Woodell, G. A., 1997b. Properties and performance of a center/surround retinex. *IEEE Transactions on Image Processing*, 6(3), pp. 451-462.
- Jobson, D., Rahman, Z.-u. and Woodell, G., 1997a. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions Image Processing*, 6(7), pp. 965 - 976.
- Jourlin, M. and Pinoli, J.-C., 1987. Logarithmic image processing. *Acta Stereologica*, Volume 6, p. 651–656.

- Jourlin, M. and Pinoli, J.-C., 1988. A model for logarithmic image processing. *Journal of Microscopy*, Volume 149, p. 21–35.
- Kam, F., 2009. *Classification Techniques for Hyperspectral Remote Sensing*, Swindon: Cranfield University CDS, Mphil thesis.
- Klinker, G. J., Shafer, S. A. and Kanade, T., 1988. The measurement of highlights in color images. *International Journal of Computer Vision*, 2(1), pp. 7-32.
- Kries, J. v., 1970. Influence of adaptation on the effects produced by luminous stimuli. In: D. L. MacAdam, ed. *Sources of Color Science*. Cambridge: The MIT Press, pp. 109-119..
- Land, E. H., 1977. The retinex theory of color vision. *Scientific American*, 237(6), p. 108–128.
- Land, E. H., 1986. An alternative technique of the computation of the designator in the Retinex theory of colour vision. *Proc. of the National Academy of Science USA*, 83(May), pp. 2078-3080.
- Land, E. H. and McCann, a. J. J., 1971. Lightness and Retinex Theory. *J. Opt. Soc. Am.*, 61(1), pp. 1-11.
- Lee, H., 1986. Method for computing the scene-illuminant from specular highlights. *Journal of the Optical Society of America*, 3(10), p. 1694–1699.
- Manolakis, D., Marden, D. and Shaw, G., 2003. Hyperspectral image processing for automatic target detection applications. *LINCOLN LABORATORY JOURNAL*, 14(1), pp. 79-116.
- McCahill, M. and Norris, C., 2003. Estimating the extent, sophistication and legality of CCTV in London. In: M. Gill, ed. *CCTV*. s.l.:Perpetuity Press, pp. 51-66.
- Mehetre, B. M., Kankanhalli, M. S., Narasimhalu, A. D. and Man, G. C., 1995. Color matching for image retrieval. *Pattern Recognition Letters*, 16(3), pp. 325-331.

- Meyer, T., 2010. *VARIATIONAL APPROACH TO RETINEX THEORY*, Los Angeles, USA: UCLA Mathematics Department.
- Mittal, A. and Davis, L. S., 2003. M 2 Tracker: a multi-view approach to segmenting and tracking people in a cluttered scene. *International Journal of Computer Vision*, 51(3), pp. 189-203.
- Orsini, G., Ramponi, G., Carrai, P. and Federico, R. D., 2003. *A modified retinex for image contrast enhancement and dynamics control*. Barcelona, Spain, IEEE Int. Conf. on Image Process.,.
- Orwell, J., Remagnino, P. and Jones, G., 1999. *Multi-camera colour tracking*. Kyoto, Japan, IEEE Workshop on Visual Surveillance.
- Panetta, K., Agaian, S., Zhou, Y. and Wharton, E., 2011. Parameterized Logarithmic Framework for Image Enhancement. *IEEE Transactions on Systems, Man, and Cybernetics*, 41(2), p. 460–473.
- Park, U., Jain, A.K., Kitahara, I., Kogure, K. and Hagita, N., 2006. *ViSE: Visual Search Engine Using Multiple Networked Cameras*. Hong Kong, China, IEEE International Conference Pattern Recognition (ICPR).
- PETS, 2007. *PETS 2007 Benchmark Data*. [Online] Available at: <http://pets2007.net> [Accessed 1 January 2014].
- Porikli, F., 2003. *Inter-camera color calibration by correlation model function*. Barcelona, IEEE, International Conference Image Processing (ICIP).
- Prosser, B., Gong, S. and Xiang, T., 2008. *Multi-camera Matching using Bi-Directional Cumulative Brightness Transfer Functions*. Leeds, British Machine Vision Conference (BMVA).
- Rahman, Z., G. A. Woodell and Jobson, D. J., 1997. *A Comparison of the Multiscale Retinex With Other Image Enhancement Techniques*,. s.l., Proceedings of the ISandT 50th Anniversary Conference.

Rahman, Z., Jobson, D. J. and Woodell, G. A., 1996. *Multiscale Retinex for Colour Rendition and Dynamic Range Compression*. Denver, SPIE Applications of Digital Image Processing.

Rahman, Z.-u., Jobson, D. J. and Woodell, G. A., 1998. *Resiliency of the Multiscale Retinex Image Enhancement Algorithm Techniques*. s.l., Proceedings of the ISandT Sixth Annual Color Conference.

Reed, I. S. and Xiaoli, Y., 1990. Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution. *IEEE Transactions Acoustics, Speech and Signal Processing*, 38(10), pp. 1760-1760.

Reinhard, E., Adhikhmin, M., Gooch, B. and Shirley, P., 2001. Color transfer between images. *IEEE Computer Graphics and Applications*, 21(5), pp. 34 - 41.

Rosenberg, C., Minka, T. and Ladsariya, A., 2003. *Bayesian Color Constancy with Non-Gaussian Models*. British Columbia, Canada, Neural Information Processing Systems (NIPS).

Sharma, G., Wu, W. and Dalal, E. N., 2005. The CIEDE2000 color-difference formula: implementation notes, supplementary test data, and mathematical observations. *Wiley Periodicals, Color Research and Application*, 30(1), pp. 21-31.

Smith, G. J. D., 2004. Behind the Screens: Examining Constructions of Deviance and Informal Practices among CCTV Control Room Operators in the UK. *Surveillance and Society*, 2(2/3), pp. 376-395.

Soori, U., Yuen, P., Han, J W., Ibrahim, I., Chen, W., Hong, K., Merfort, C., James, D. and Richardson, M., 2013. Target recognitions in multiple-camera closed-circuit television using color constancy. *SPIE Optical Engineering*, 52(4), pp. 047202-1- 047202-13.

Soori, U., Yuen, P. W. T. and Ibrahim, I., 2011. *Colour invariant target recognition in multiple camera CCTV surveillance*. Prague, SPIE.

- Stein, D. W. J., Beaven, S.G., Hoff, L.E., Winter, E.M., Schaum, AP. and Stocker, AD, 2002. Anomaly detection from hyperspectral imagery. *IEEE Signal Processing Magazine, IEEE*, 19(1), pp. 58-69.
- Sun, B., Chen, W., Li, H., Tao, W. and Li, J., 2007. *Modified Luminance Based Adaptive MSR*. Sichuan, IEEE International Conference on Image and Graphics.
- Sun, B., Tao, W., Chen, W., Li, H., Li, J. and Guo, W., 2008. *Luminance Based MSR for Color Image Enhancement*. SANYA, China, IEEE Image and Signal Processing.
- Swain, M. and Ballard, D., 1990. *Indexing via color histograms*. Osaka, Third International Conference on Computer Vision.
- Tan, R. T., Nishino, K. and Ikeuchi, K., 2004. Color Constancy through Inverse-Intensity Chromaticity Space. *Optical Society of America*, 21(3), p. 321–334.
- Tao, L. and Asari, a. V. K., 2003. *Modified Luminance Based MSR for Fast and Efficient Image Enhancement*. Washington, USA, IEEE Applied Imagery Pattern Recognition Workshop.
- Tao, L. and Asari, V. K., 2005. Adaptive and integrated neighborhood-dependent approach for nonlinear enhancement of color images. *SPIE Journal of Electronic Imaging*, 14(4), pp. 043006-1 - 043006-14.
- Tao, L., Seow, M.-J. and Asari, V. K., 2006. *Nonlinear image enhancement to improve face detection in complex lighting environment*. San Jose, USA, SPIE Proc. of Image Processing: Algorithms and Systems, Neural Networks, and Machine.
- Tao, L., Tompkins, R. and Asari, V., 2005. *An illuminance-reflectance nonlinear video enhancement model for homeland security applications*. Washington, USA, IEEE Applied Imagery and Pattern Recognition Workshop.
- Theodoridis, S., Pikrakis, A., Koutroumbas, K. and Cavouras, D., 2010. Classifiers based on Bayes decision theory. In: *Introduction to Pattern*

Recognition: A Matlab Approach . Burlington, Massachusetts: Academic press, p. 1–27.

Tipping, M. and Bishop, C., 1999. Probabilistic principal component analysis. *Journal of the Royal Statistical Society, Series B*, 11(Part 3), p. 611–622.

Tsai, C.-Y., 2012. A Fast Dynamic Range Compression With Local Contrast Preservation Algorithm and Its Application to Real-Time Video Enhancement. *IEEE TRANSACTIONS ON MULTIMEDIA*, 14(4), pp. 1140 - 1152.

Tsai, C.-Y. and Chou:, C.-H., 2011. A novel simultaneous dynamic range compression and local contrast enhancement algorithm for digital video cameras. *Springer Journal on Image and Video Processing*, 2011(6), pp. 1-19.

UK Home Office, 2008. *Imagery library for intelligent detection systems (iLIDS) multiple camera tracking scenario definition*. [Online] Available at: www.homeoffice.gov.uk/science-research/hosdb/i-lids/ [Accessed 1 January 2010].

Unaldi, N., Asari, V. K. and Rahman, Z.-u., 2009. *Fast and robust wavelet-based dynamic range compression and contrast enhancement model with color restoration*. Orlando, USA, SPIE Visual Information Processing.

Wang, X., Doretto, G., Sebastian, T., Rittscher, J., Tu, P, 2007. *Shape and Appearance Context Modeling*. Rio de Janeiro, Brazil, IEEE International Conference Computer Vision (ICCV).

Weijer, J. v. d., Gevers, T. and Gijssenij, A., 2007. Edge-Based Color Constancy. *IEEE Trans. Image Process.*, 16(9), pp. 2207 - 2214.

West, J. E., 2005. *Matched Filter Stochastic Background Characterization for Hyperspectral Target Detection*, s.l.: Rochester Institute of Technology .

Zhang, R., Huang, Y. and Zhao, Z., 2011. *A Ultrasound Liver Image Enhancement Algorithm Based on Multi-Scale Retinex Theory*. Wuhan, IEEE International Conference Bioinformatics and Biomedical Engineering.

Zhu, J., Liao, S., Lei, Z., Yi and Dong L., S.Z., 2013. *Pedestrian Attribute Classification in Surveillance: Database and Evaluation*. Sydney, NSW, IEEE International Conference on Computer Vision Workshops (ICCVW).

APPENDICES

Appendix A : Limitations of CC algorithms for target detections

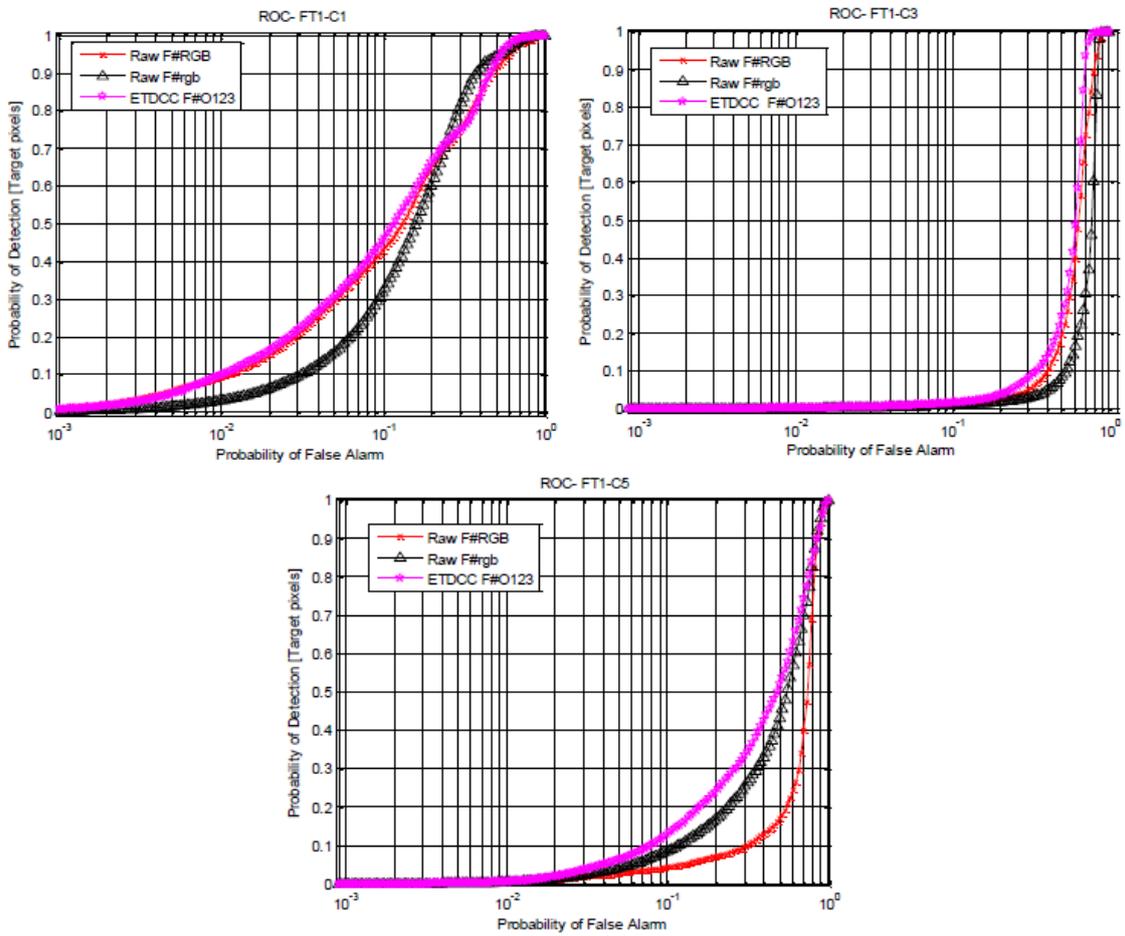
During the course of this work it is found that some targets, especially those in weak colours such as grey, white and black etc., are difficult to recognise/detect using the spectral feature for the detection. Figure A-1 depicts 3 examples of these targets extracted from the iLIDS data set with clothing in white, grey and black colours and they are labelled as targets FT1, FT2 and FT3 respectively. The detection performances of these targets are shown in Figure A-2 (a)-(c) which indicates high false alarm for both the data using the RAW RGB and after transformed by the ETDCC algorithm. It is surprised to observe a worse detection for the data after the CC treatment, even though the colour of the target have been maintained to a constant as depicted by the ΔE_{ab} assessment over the 10 video clips as shown in Figure A-3(b).

The large false alarm rate is due to the background which contains many objects in white/grey/black colour. It is well-known that the discrimination ability of the broadband RGB is much weaker than that of the narrow band hyperspectral imaging (HSI). It is believed that the detection of these weak colour targets may be improved if narrow spectral bands were employed.

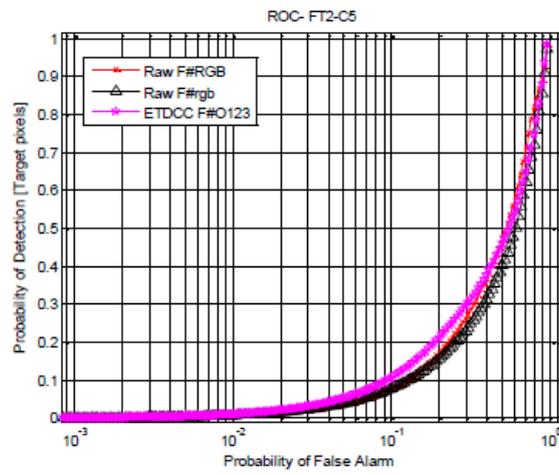
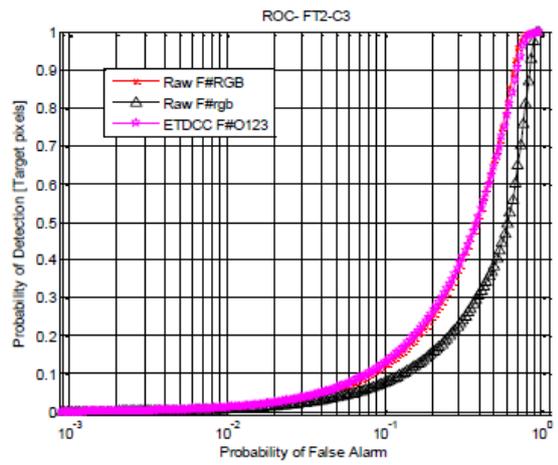
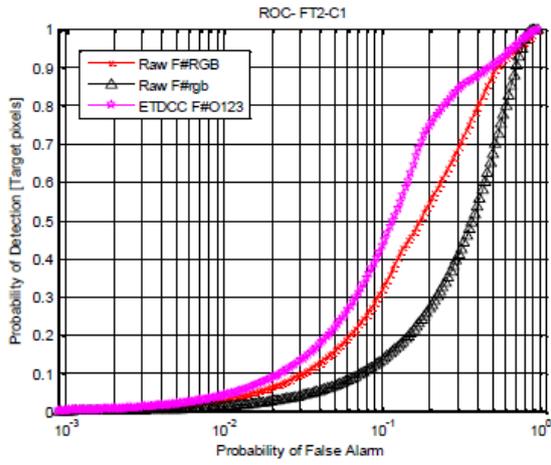
The 'worse' detection result observed after CC treatment is thought may be due to the correction of these weak colours close to their intrinsic values which are highly resembled to many clutters in the background. This could be one of the reasons to induce a worse detection as seen in the experiment.



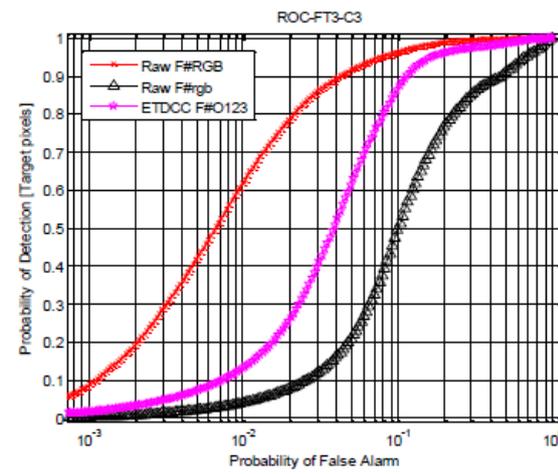
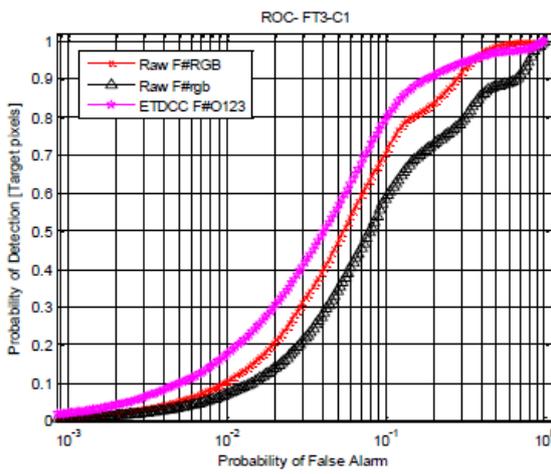
Figure A-1: Shows examples of targets that are difficult to detect using spectral feature: left to right FT1 (white), FT2 (grey) and FT3 (black).

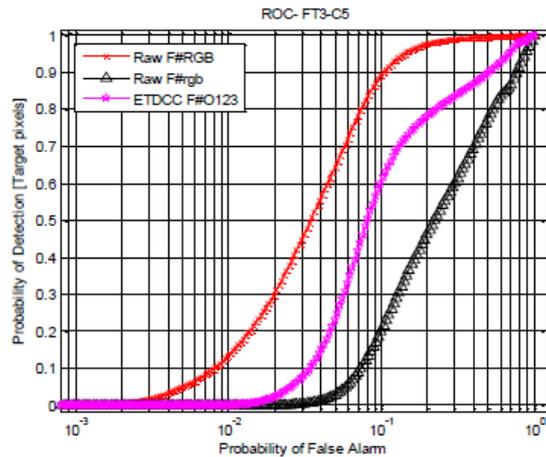


(a)



(b)



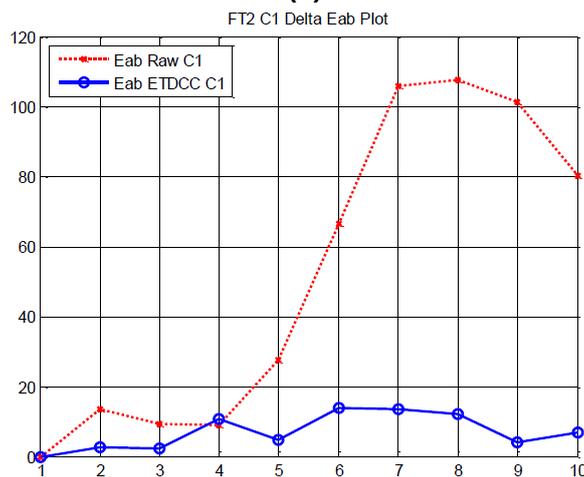


(c)

Figure A-2: Shows the ROC results for the detection of targets in 3 camera views of: (a) FT1 (white), (b) FT2 (grey) and (c) FT3 (black).



(a)



(b)

Figure A-3: (a) top row: raw colour of FT2 (grey) for the 10 frames of the video clip. Bottom row: same data after transformed by the ETDCC algorithm. (b) Shows the ΔE_{ab} assessment of the data shown in (a).

Appendix B : Variable C in the PWNLTF

The pixel wise non-linear function as depicted in Equation (5.8) consists of a free parameter C which sets the form of the transfer function. This parameter C has been chosen by experiment. Figure B-1 shows the detection result of target T1 for using various C values in the PWNLTF which is a part of the ETDCC algorithm. It is seen that a moderate value of $C \sim 0.2$ is more appropriate to this data set and this setting of C have been used for all detection results reported in this thesis.

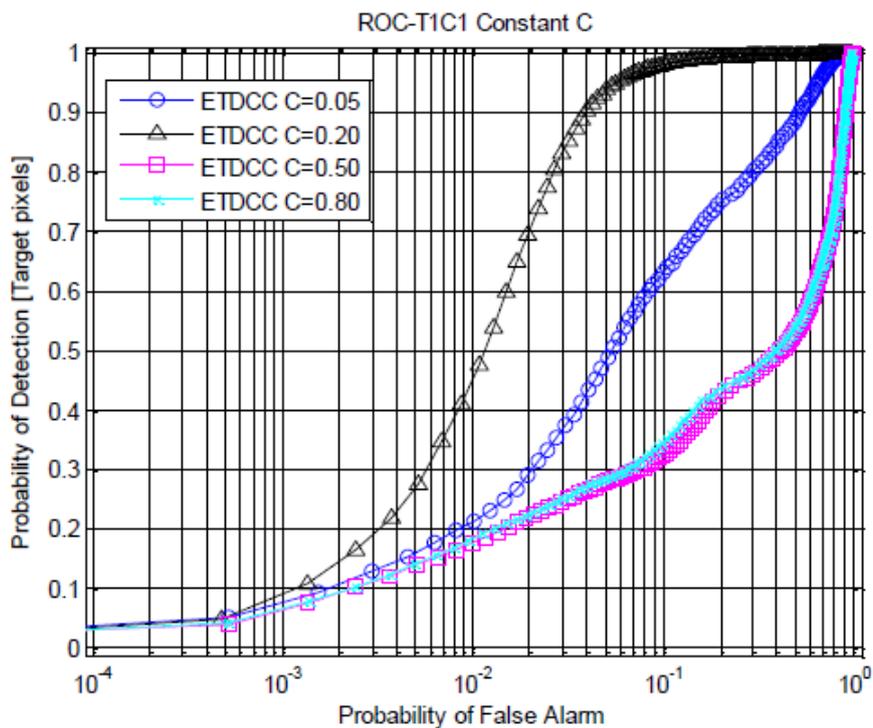


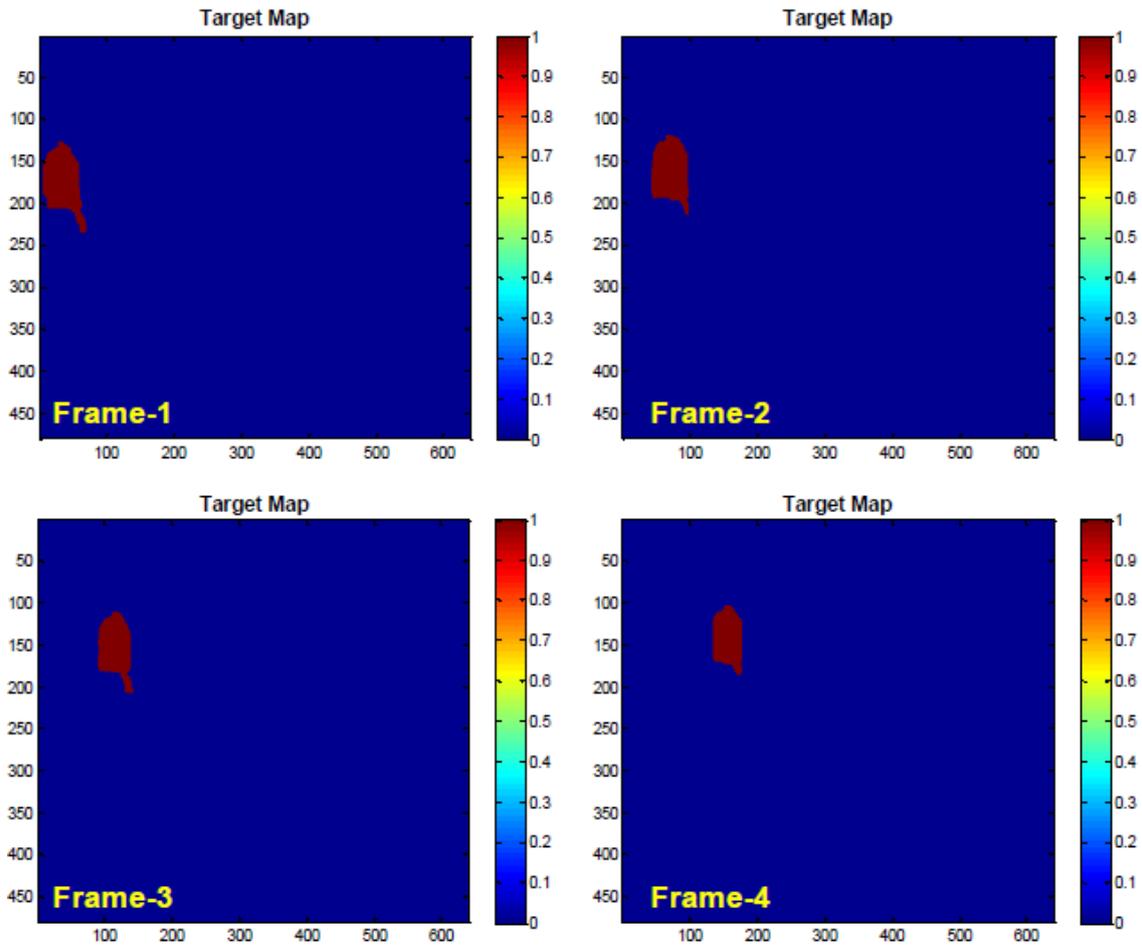
Figure B-1: Show the various C values for the detection of T1 after the data is processed by PWNLTF using various C values. $C=0.2$ has been chosen for all the detection presented in this work.

Appendix C : Frame to frame target detections

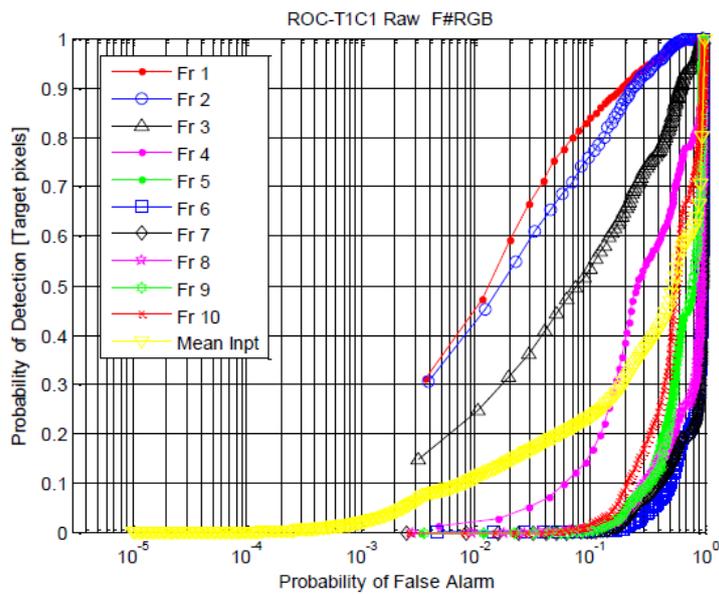
The process for the detection of targets from video clips is elucidated in more details in this appendix. As mentioned in section 3.4.1 that each target is detected from 10 frames of a video clip and the frame to frame detection results are shown in Figure C-1.

Figure C-1(a) shows the first four frames of the ground truth target map for the Target 1 as an example. Figure C-1(b & C) shows the detection ROC results of the raw data & the images after processed by ETDCG respectively. The MD classifier has been utilised for the detection. Note that the target is almost missed completely after a few video frames in the raw data set, with a very inconsistent ROC scatter over the 10 frame of detections. This is very different in the ETDCG treated data (Figure C-1(c)) which exhibits relatively more consistent ROC over the 10 detections. Figure C-1(d) highlights the detection maps of the first 4 frame of detections for the raw data set at a constant PFA of 0.02. The correctly identified target pixels are labelled in brown in the figure. It is quite clear that the target is almost missed at the 4th frame of detection. Figure C-1(e) shows the detection maps of the first 4 frames of images after they are treated by the ETDCG algorithm: the target is seen more faithfully detected.

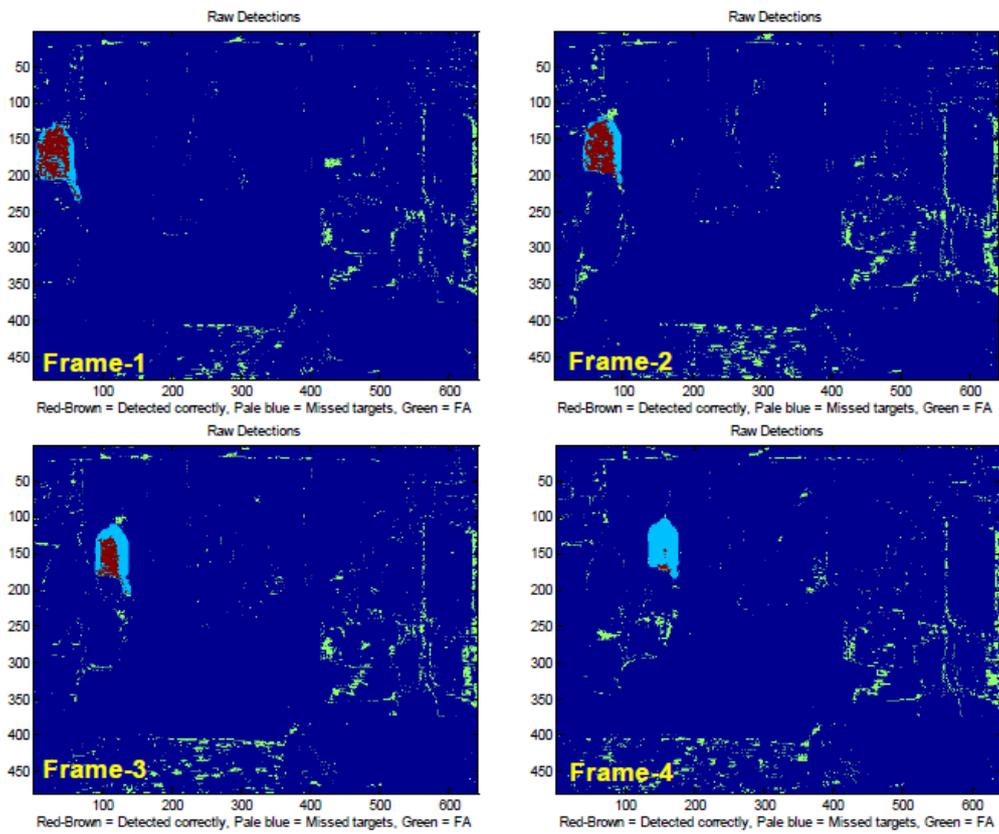
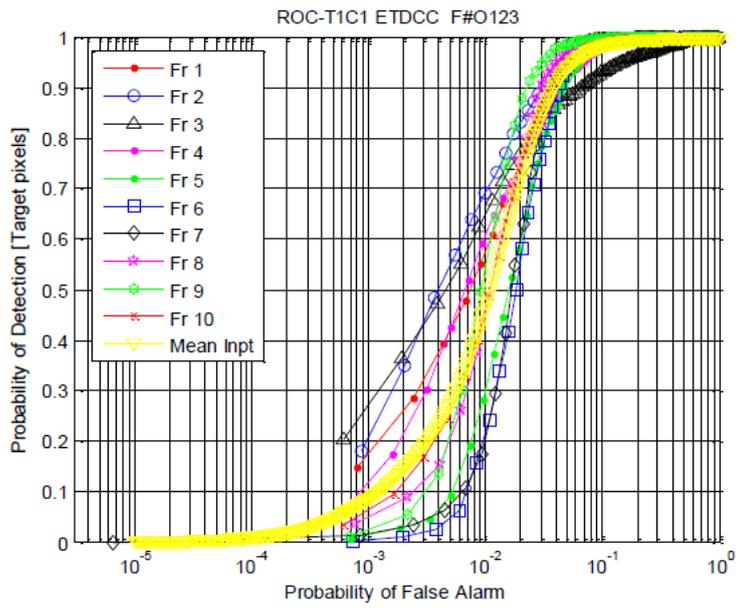
Note that the average of these ROC as depicted in yellow plots of Figure C-1(b & C) represent the representative detection results of the target for the raw and CC treated data respectively.

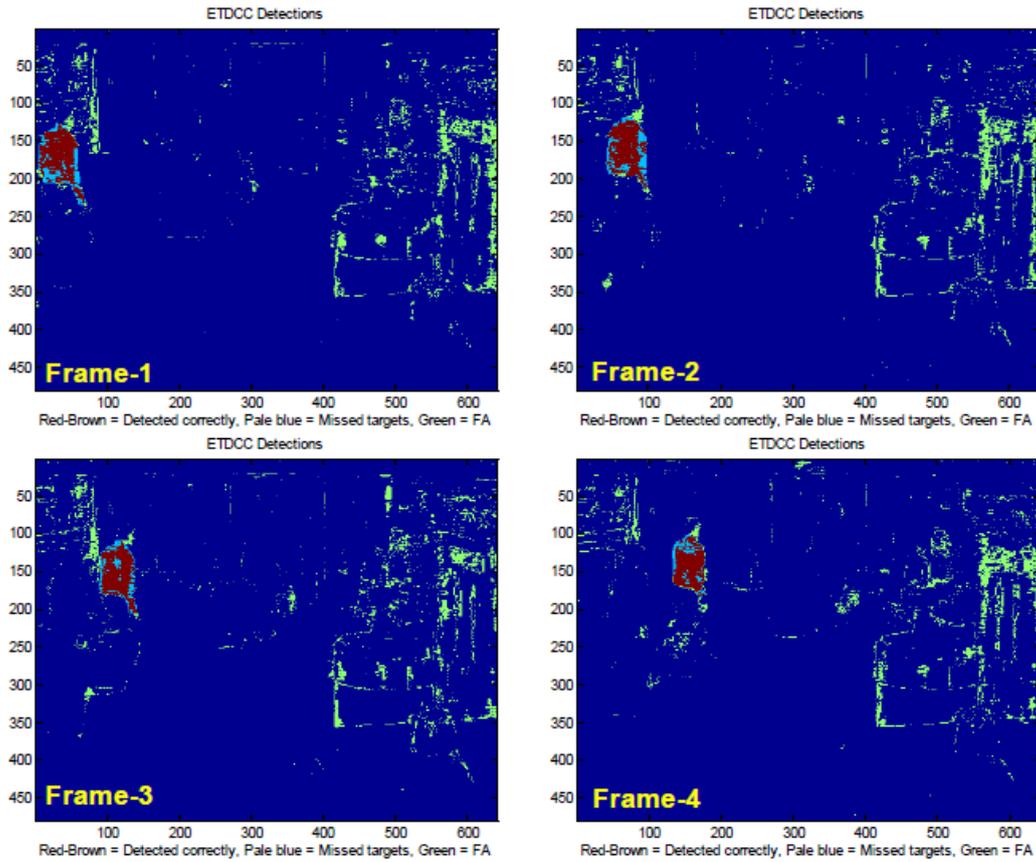


(a)



(b)





(e)

Figure C-1: Show an example of the frame-to-frame detection process employed in this work (a) Ground truth target map for the first 4 frames of images (b) ROC detections for the 10 frames of the raw data (c) ROC detections for the 10 frames of the data after treated by the ETDCC algorithm, (d) Detection map of the first 4 frames of data using the raw data as test images at PFA 0.02, (e) Detection map of the first 4 frames of data after they are treated by the ETDCC at PFA 0.02.